```
                     GenCore version 4.5
          Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          June 14, 2001, 08:04:29 ; Search time 31.39 Seconds
                                           (without alignments)
                                           2016.320 Million cell updates/sec

Title:           US-09-445-223-1
Perfect score:   2829
Sequence:        1 MNGEAICSALPTIPYHKLAD..........PEILVVSRSPSLNLLQNKSM 540

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters:      374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_15:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_unclassified:*
                13:  sp_vertebrate:*
                14:  sp_virus:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                   %
Result           Query
  No.   Score    Match Length DB  ID                      Description
 -------------------------------------------------------------------------------
    1    2823    99.8    540  4  O43353                  O43353 homo sapien
    2     401    14.2    486 11   Q9QZL0                  Q9qzl0 mus musculu
    3   389.5    13.8    806 10   Q9ZSD8                  Q9zsd8 lycopersico
    4   389.5    13.8    829 10   Q9ZSD9                  Q9zsd9 lycopersico
    5   385.5    13.6    829 10   O24027                  O24027 lycopersico
    6     375    13.3    518  4  Q9Y572                  Q9y572 homo sapien
    7   359.5    12.7    963 10   Q9STU7                  Q9stu7 arabidopsis
    8     353    12.5    773 10   Q9M8C1                  Q9m8c1 arabidopsis
    9   348.5    12.3    370 10   Q9S7D5                  Q9s7d5 arabidopsis
   10   348.5    12.3    763 10   Q9M8C3                  Q9m8c3 arabidopsis
   11     340    12.0    981 10   O65833                  O65833 lycopersico
   12   337.5    11.9   1030 10   Q9SFN9                  Q9sfn9 arabidopsis
   13   335.5    11.9    736 10   O82754                  O82754 arabidopsis
   14   334.5    11.8    835  4  Q9Y2V6                  Q9y2v6 homo sapien
   15   332.5    11.8    405 10   Q9LVQ9                  Q9lvq9 arabidopsis
   16     329    11.6    671 10   Q9M8C2                  Q9m8c2 arabidopsis
   17   327.5    11.6    406 10   O23719                  O23719 arabidopsis
   18     327    11.6    678  5  Q9V3Q6                  Q9v3q6 drosophila
   19   317.5    11.2    462 10   Q39886                  Q39886 glycine max
   20   316.5    11.2    776 10  O80963                  O80963 arabidopsis
   21   313.5    11.1    888 11  Q63796                  Q63796 rattus norv
   22   313.5    11.1   1200 10  Q9MAE9                  Q9mae9 arabidopsis
   23     312    11.0    579  4 O43317                  O43317 homo sapien
   24     311    11.0    579 11  Q62073                  Q62073 mus musculu
   25     310    11.0    546 10  O22558                  O22558 arabidopsis
   26     309    10.9    438 10  Q9XI87                  Q9xi87 arabidopsis
   27     309    10.9    445 10  Q9LMF8                  Q9lmf8 arabidopsis
   28     308    10.9    491  4  Q9NZ70                 Q9nz70 homo sapien
   29     308    10.9    539  4  Q9NTR1                 Q9ntr1 homo sapien
   30     307    10.9    553 10  O81808                  O81808 arabidopsis
   31   306.5    10.8    475 10  Q9STG5                  Q9stg5 arabidopsis
   32   306.5    10.8    888 11  Q60700                  Q60700 mus musculu
   33   306.5    10.8    988 10  Q9LM32                  Q9lm32 arabidopsis
   34   304.5    10.8    616 13  O73613                  O73613 xenopus lae
   35   304.5    10.8    888 11  P70286                  P70286 mus musculu
   36     304    10.7    451  4  Q9NTR3                 Q9ntr3 homo sapien
   37   301.5    10.7    886 10  Q9LYI8                  Q9lyi8 arabidopsis
   38   300.5    10.6    606  4  O43318                 O43318 homo sapien
   39     300    10.6    443 10  O65239                  O65239 arabidopsis
   40   299.5    10.6   1055 13  O42440                  O42440 oryzias lat
   41     299    10.6    567 10  Q9ZQN5                  Q9zqn5 arabidopsis
   42     298    10.5    518  4  O43319                 O43319 homo sapien
   43   297.5    10.5    545 10  Q9SZM7                  Q9szm7 arabidopsis
   44     297    10.5   1055 13  Q9PVW4                  Q9pvw4 oryzias lat
   45   296.5    10.5    566  4  Q9NTR2                 Q9ntr2 homo sapien

                              ALIGNMENTS

RESULT   1
O43353
ID   O43353     PRELIMINARY;      PRT;   540 AA.
AC   O43353;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   SERINE/THREONINE KINASE RICK.
GN   RICK OR RIP2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98241596; PubMed=9575181;
RA   Inohara N., del Peso L., Koseki T., Chen S., Nunez G.;
RT   "RICK, a novel protein kinase containing a caspase recruitment domain,
RT   interacts with CLARP and regulates CD95-mediated apoptosis.";
RL   J. Biol. Chem. 273:12296-12300(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Ozersky P., Holmes A., Broy M.;
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98307936; PubMed=9642260;
RA   McCarthy J.V., Ni J., Dixit V.M.;
RT   "RIP2 is a novel NF-kappaB-activating and cell death-inducing
RT   kinase.";
RL   J. Biol. Chem. 273:16968-16975(1998).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Thome M., Hofmann K., Burns K., Martinon F., Bodmer J.-L.,
RA   Mattmann C., Tschopp J.;
RT   "Identification of CARDIAK, a RIP-like kinase that associates with
RT   caspase-1.";
RL   Curr. Biol. 8:0-0(1998).
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Platzer M., Varon R.;
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; AF027706; AAC34970.1; -.
DR   EMBL; AC004003; AAC24561.1; -.
DR   EMBL; AF078530; AAC27722.1; -.
DR   EMBL; AF064824; AAC25668.1; -.
DR   EMBL; AF117829; AAD04634.1; -.
DR   HSSP; P00523; 2PTK.
DR   INTERPRO; IPR000719; -.
DR   INTERPRO; IPR001315; -.
DR   INTERPRO; IPR002290; -.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00619; CARD; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Kinase.
SQ   SEQUENCE   540 AA;  61194 MW;  575A692239505792 CRC64;

  Query Match             99.8%;  Score 2823;  DB 4;  Length 540;
  Best Local Similarity   99.8%;  Pred. No. 1.9e-220;
  Matches  539;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1 MNGEAICSALPTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSER 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 MNGEAICSALPTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSER 60

Qy       61 KDVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPL 120
            |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db       61 KDVLREAEILHKARFSYILPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPL 120

Qy      121 RFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRS 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 RFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRS 180

Qy      181 SKSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMY 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 SKSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMY 240

Qy      241 SVSQGHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEI 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 SVSQGHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEI 300

Qy      301 TFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLNIPVNHGPQEESCGSSQLHENSGSPET 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 TFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLNIPVNHGPQEESCGSSQLHENSGSPET 360

Qy      361 SRSLPAPQDNDFLSRKAQDCYFMKLHHCPGNHSWDSTISGSQRAAFCDHKTTPCSSAIIN 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 SRSLPAPQDNDFLSRKAQDCYFMKLHHCPGNHSWDSTISGSQRAAFCDHKTTPCSSAIIN 420

Qy      421 PLSTAGNSERLQPGIAQQWIQSKREDIVNQMTEACLNQSLDALLSRDLIMKEDYELVSTK 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 PLSTAGNSERLQPGIAQQWIQSKREDIVNQMTEACLNQSLDALLSRDLIMKEDYELVSTK 480

Qy      481 PTRTSKVRQLLDTTDIQGEEFAKVIVQKLKDNKQMGLQPYPEILVVSRSPSLNLLQNKSM 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 PTRTSKVRQLLDTTDIQGEEFAKVIVQKLKDNKQMGLQPYPEILVVSRSPSLNLLQNKSM 540

RESULT   2
Q9QZL0
ID   Q9QZL0      PRELIMINARY;      PRT;   486 AA.
AC   Q9QZL0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   RECEPTOR INTERACTING PROTEIN 3.
GN   RIP3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RA   Pazdernik N.J., Donner D.B., Goebl M.G., Harrington M.A.;
RT   "Mouse Interacting Protein 3 Does Not Contain a Caspase-Recruiting or
RT   a Death Domain but Induces Apoptosis and Activates NF-kappaB.";
RL   Mol. Cell. Biol. 19:0-0(1999).
DR   EMBL; AF178953; AAF03133.1; -.
DR   INTERPRO; IPR000719; -.
DR   INTERPRO; IPR001245; -.
DR   INTERPRO; IPR002290; -.
DR   PFAM; PF00069; pkinase; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Receptor.
SQ   SEQUENCE   486 AA;  53336 MW;  DD264E69187D3436 CRC64;

  Query Match             14.2%;  Score 401;  DB 11;  Length 486;
  Best Local Similarity   28.0%;  Pred. No. 2.6e-24;
  Matches  140;  Conservative   70;  Mismatches  192;  Indels   98;  Gaps   18;

Qy        8 SALPTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSERKDVLREA 67
            ||:| :   :|  | :: :|  | |  | |  |   |||| ::          | :  |
Db       12 SAVPLVSREELKKLEFVGKGGFGVVFRAHHRTWNHDVAVKIVN---------SKKISWEV 62

Qy       68 EILHKARFSYIFPILGICNEPEF-----LGIVTEYMPNGSLNELLHRKTEYPDVAWPLRF 122
            : :   |   :  :||:  : ::       :|| :| ||||  ||  :   |   |||
Db       63 KAMVNLRNENVLLLLGVTEDLQWDFVSGQALVTRFMENGSLAGLLQPECPRP---WPLLC 119

Qy      123 RILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRS-S 181
            |:| |: ||: |||:: ||||| |||  ||||| | | |:|||||| ::  | | | | |
Db      120 RLLQEVVLGMCYLHSLDPPLLHRDLKPSNILLDPELHAKLADFGLSTFQGGSQSGSGSGS 179

Qy      182 KSAPEGGTIIYMPPE-----NYEPGQKSRASIKHDIYSYAVITWEVLS-RKQPFEDVTNP 235
             |   |||: |: ||     |       :||   |:||: :: | ||: |:    |  |:
Db      180 GSRDSGGTLAYLDPELLFKVNL------KASKASDVYSFGILVWAVLAGREAELVDKTSL 233

Qy      236 LQIMYSVSQGHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLR 295
            ::      |   |:        : |   ::  |:   |    : ||||  |    ||  :
Db      234 IRETVCDRQSRPPLTELPPGSPETPGLEKLKELMIHCWGSQSENRPSFQDC----EP--K 287

Qy      296 TFEEITFLE-----AVIQLKKTKLQSVSSAIHL-------------CDKK----KMELSL 333
            | |    ::     || ::|    |  ||  :|             | ::    ||   |
Db      288 TNEVYNLVKDKVDAAVSEVKHYLSQHRSSGRNLSAREPSQRGTEMDCPRETMVSKMLDRL 347

Qy      334 NIPVNHGPQEESCGSSQLHENSGSPETSRSLPAPQDNDFLSRKAQDCYFMKLH------- 386
            ::    ||    |   |  : |  | |    ||   :| ::   |  : :
Db      348 HLEEPSGPVPGKCPERQAQDTSVGPAT----PARTSSDPVAGTPQIPHTLPFRGTTPGPV 403

Qy      387 --HCPGNHSWDSTISGSQRAAFCDHKTTP--------------CS--------SAIINPL 422
                || |   :   |  |     :  ||              ||        | :  |
Db      404 FTETPGPHPQRN--QGDGRHGTPWYPWTPPNPMTGPPALVFNNCSEVQIGNYNSLVAPPR 461

Qy      423 STAGNS---ERLQPGIAQQW 439
            :|| :|   :: | |  : |
Db      462 TTASSSAKYDQAQFGRGRGW 481

RESULT   3
Q9ZSD8
ID   Q9ZSD8      PRELIMINARY;      PRT;   806 AA.
AC   Q9ZSD8;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   ETHYLENE-INDUCIBLE CTR1-LIKE PROTEIN KINASE.
OS   Lycopersicon esculentum (Tomato).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC   Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4081;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=CV. AILSA CRAIG;
RA   Kannan P., Giovannoni J.J.;
RT   "Ethylene-inducible tomato CTR1-like protein kinase.";
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF110519; AAD10057.1; -.
DR   HSSP; P06213; 1IRK.
DR   INTERPRO; IPR000719; -.
DR   INTERPRO; IPR001245; -.
DR   INTERPRO; IPR002290; -.
DR   PFAM; PF00069; pkinase; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Kinase.
SQ   SEQUENCE   806 AA;  89387 MW;  9102529F49549A61 CRC64;

  Query Match             13.8%;  Score 389.5;  DB 10;  Length 806;
  Best Local Similarity   34.0%;  Pred. No. 4.5e-23;
  Matches  107;  Conservative   55;  Mismatches  104;  Indels   49;  Gaps   14;

Qy      3 GEAICSALPT-----------IPYHKLADLRYLSRGASGTVSSARHADWR-VQVAVKHL- 49
          |  : ||:|:           ||:: |  :  :  |: |||      ||    |||| |
Db    506 GGHVVSAIPSEELDLDVEEFNIPWNDLVLMEKIGAGSFGTV---HRGDWHGSDVAVKILM 562

Qy     50 --HIHTPLLDSERKDVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELL 107
              |   |    |: |||  |: : |   |   :|   :|  | |||||:  |||  ||
Db    563 EQDFHAERL----KEFLREVAIMKRLRHPNIVLFMGAVIQPPNLSIVTEYLSRGSLYRLL 618

Qy    108 HR--KTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADF 165
          |:    :  |   ||   : :::| |:||||   ||::| |||: |:|:| :: ||| ||
Db    619 HKPGAKKVLDERRPL--CLAYDVANGMNYLHKRNPPIVHRDLKSPNLLVDKKYTVKICDF 676

Qy    166 GLSKWRMMSLSQSRSSKSAPEGGTIIYMPPE--NYEPGQKSRASIKHDIYSYAVITWEVL 223
          |||:::  :   |:::   ||     :| ||    ||  :     | |:||: || ||:
Db    677 GLSRFKANTFLSSKTAAGTPE-----WMAPEVIRDEPSNE-----KSDVYSFGVILWELA 726

Qy    224 SRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIPH--RARMISLIESGWAQNPDERP 281
          : :||:  :  || |::  :|        |  : |  |||     |:  :||: ||  | :||
Db    727 TLQQPWNKL-NPPQVIAAVG------FNRKKL--DIPSVLNPRVAIIIEACWANEPWKRP 777

Qy    282 SFLKCLIELEPVLRT 296
          ||   :  | | |::
Db    778 SFSTIMDMLRPHLKS 792


RESULT   4
Q9ZSD9
ID   Q9ZSD9      PRELIMINARY;      PRT;   829 AA.
AC   Q9ZSD9;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   ETHYLENE-INDUCIBLE CTR1-LIKE PROTEIN KINASE.
OS   Lycopersicon esculentum (Tomato).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC   Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4081;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. AILSA CRAIG;
RA   Kannan P., Giovannoni J.J.;
RT   "Ethylene-inducible tomato CTR1-like protein kinase.";
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF110518; AAD10056.1; -.
DR   HSSP; P06213; 1IRK.
DR   INTERPRO; IPR000719; -.
DR   INTERPRO; IPR001245; -.
DR   INTERPRO; IPR002290; -.
DR   PFAM; PF00069; pkinase; 1.
DR   PRINTS; PR00109; TYRKINASE.
```

```
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Kinase.
SQ   SEQUENCE   829 AA;  91889 MW;  9E976BF7640CE11D CRC64;

  Query Match             13.8%;  Score 389.5;  DB 10;  Length 829;
  Best Local Similarity   34.0%;  Pred. No. 4.6e-23;
  Matches  107;  Conservative   55;  Mismatches  104;  Indels   49;  Gaps   14;

Qy      3 GEAICSALPT-----------IPYHKLADLRYLSRGASGTVSSARHADWR-VQVAVKHL- 49
          |  : ||:|:           ||:: |  :  :  |: |||      ||    |||| |
Db    529 GGHVVSAIPSEELDLDVEEFNIPWNDLVLMEKIGAGSFGTV---HRGDWHGSDVAVKILM 585

Qy     50 --HIHTPLLDSERKDVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELL 107
              |   |    |: |||  |: : |   |   :|   :|  | |||||:  |||  ||
Db    586 EQDFHAERL----KEFLREVAIMKRLRHPNIVLFMGAVIQPPNLSIVTEYLSRGSLYRLL 641

Qy    108 HR--KTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADF 165
          |:    :  |   ||   : :::| |:||||   ||::| |||: |:|:| :: ||| ||
Db    642 HKPGAKKVLDERRPL--CLAYDVANGMNYLHKRNPPIVHRDLKSPNLLVDKKYTVKICDF 699

Qy    166 GLSKWRMMSLSQSRSSKSAPEGGTIIYMPPE--NYEPGQKSRASIKHDIYSYAVITWEVL 223
          |||:::  :   |:::   ||     :| ||    ||  :     | |:||: || ||:
Db    700 GLSRFKANTFLSSKTAAGTPE-----WMAPEVIRDEPSNE-----KSDVYSFGVILWELA 749

Qy    224 SRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIPH--RARMISLIESGWAQNPDERP 281
          : :||:  :  || |::  :|        |  : |  |||     |:  :||: ||  | :||
Db    750 TLQQPWNKL-NPPQVIAAVG------FNRKKL--DIPSVLNPRVAIIIEACWANEPWKRP 800

Qy    282 SFLKCLIELEPVLRT 296
          ||   :  | | |::
Db    801 SFSTIMDMLRPHLKS 815


RESULT   5
O24027
ID   O24027      PRELIMINARY;      PRT;   829 AA.
AC   O24027;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   HYPOTHETICAL 91.9 KDA PROTEIN.
GN   ER50.
OS   Lycopersicon esculentum (Tomato).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC   Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4081;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=UC82B;
RA   Wang Y., Li N.;
RL   Plant Physiol. 114:1135-1135(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. EVITA; TISSUE=GREEN FRUIT;
RA   Zegzouti H., Jones B., Frasse P., Marty C., Maitre B., Latche A.,
RA   Pech J.C., Bouzayen M.;
RT   "Ethylene-regulated gene expression in tomato fruit: characterization
RT   of novel ethylene-responsive and ripening-related genes isolated by
RT   differential display.";
RL   Plant J. 18:589-600(1999).
DR   EMBL; Y13273; CAA73722.1; -.
DR   EMBL; AF096250; AAD46406.1; -.
DR   HSSP; P00523; 2PTK.
DR   INTERPRO; IPR000719; -.
DR   INTERPRO; IPR002290; -.
DR   PFAM; PF00069; pkinase; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
```

```
KW   Hypothetical protein; Kinase.
SQ   SEQUENCE   829 AA;  91912 MW;  08FCF7468993537D CRC64;

  Query Match             13.6%;  Score 385.5;  DB 10;  Length 829;
  Best Local Similarity   33.2%;  Pred. No. 9.8e-23;
  Matches  104;  Conservative   56;  Mismatches  108;  Indels   45;  Gaps   12;

Qy        3 GEAICSALPT-----------IPYHKLADLRYLSRGASGTVSSARHADWR-VQVAVKHL- 49
            |  : ||:|:           ||:: |  :  :  |: |||      ||    |||| |
Db      529 GGHVVSAIPSEELDLDVEEFNIPWNDLILMEKIGAGSFGTV---HRGDWHGSDVAVKILM 585

Qy       50 --HIHTPLLDSERKDVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELL 107
                |   |    |: |||  |: : |   |   :|   :|  | |||||:  |||  ||
Db      586 EQDFHAERL----KEFLREVAIMKRLRHPNIVLFMGAVIQPPNLSIVTEYLSRGSLYRLL 641

Qy      108 HRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGL 167
            |:      :    |  : :::| |:||||   ||::| |||: |:|:| :: ||| ||||
Db      642 HKPGAREVLDERRRLCMAYDVANGMNYLHKRNPPIVHRDLKSPNLLVDKKYTVKICDFGL 701

Qy      168 SKWRMMSLSQSRSSKSAPEGGTIIYMPPE--NYEPGQKSRASIKHDIYSYAVITWEVLSR 225
            |:::  :   |:::   ||     :| ||    ||  :     | |:||: || ||: :
Db      702 SRFKANTFLSSKTAAGTPE-----WMAPEVIRDEPSNE-----KSDVYSFGVILWELATL 751

Qy      226 KQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIPH--RARMISLIESGWAQNPDERPSF 283
            :||:  : || |:: :|        | : |  |||     ::  :||: ||  | :||||
Db      752 QQPWNKL-NPPQVIAAVG------FNRKRL--DIPSDLNPQVAIIIEACWANEPWKRPSF 802

Qy      284 LKCLIELEPVLRT 296
               :  | | |::
Db      803 STIMDMLRPHLKS 815

RESULT   6
Q9Y572
ID   Q9Y572      PRELIMINARY;      PRT;   518 AA.
AC   Q9Y572;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   RIP-LIKE KINASE.
GN   RIP3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99272740; PubMed=10339433;
RA   Yu P.W., Huang B.C., Shen M., Quast J., Chan E., Xu X., Nolan G.P.,
RA   Payan D.G., Luo Y.;
RT   "Identification of RIP3, a RIP-like kinase that activates apoptosis
RT   and NFkappaB.";
RL   Curr. Biol. 9:539-542(1999).
DR   EMBL; AF156884; AAD39005.1; -.
DR   HSSP; P08631; 1AD5.
DR   INTERPRO; IPR000719; -.
DR   INTERPRO; IPR001245; -.
DR   PFAM; PF00069; pkinase; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Kinase.
SQ   SEQUENCE   518 AA;  56901 MW;  38A3ECFBEBBD4151 CRC64;

  Query Match             13.3%;  Score 375;  DB 4;  Length 518;
  Best Local Similarity   31.3%;  Pred. No. 3.6e-22;
  Matches  100;  Conservative   62;  Mismatches  127;  Indels   30;  Gaps   11;

Qy       11 PTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSERKDVLREAEIL 70
            | :   :| :   : :   |||  |:|  |   |||| ::          | : || : :
Db       14 PLVSIEELENQELVGKDGFGTVFRAQHRKWGYDVAVKIVN---------SKAISREVKAM 64
```

```
Qy       71 HKARFSYIFPILGICNEPEF-----LGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRIL 125
                  ::  : |:  :  :       :||::| ||||: ||  :   |   |||  |:|
Db       65 ASLDNEFVLRLEGVIEKVNWDQDPKPALVTKFMENGSLSGLLQSQCPRP---WPLLCRLL 121

Qy      126 HEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAP 185
             |: ||: |||:  | ||| |||  |:| | | |||:|||||| ::    ||| :    |
Db      122 KEVVLGMFYLHDQNPVLLHRDLKPSNVLPDPELHVKLADFGLSTFQ--GGSQSGTGSGEP 179

Qy      186 EGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSV-SQ 244
             |||: |: || :      :||   |:||: :: | ||: ::  |  | |  :  :| ::
Db      180 -GGTLGYLAPELF-VNVNRKASTASDVYSFGILMWAVLAGRE-VELPTEPSLVYEAVCNR 236

Qy      245 GHRPVINEESLPY---DIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEIT 301
             :|| : |  ||    : |    :  |::  |:  | :|||| :|| : : | :  |
Db      237 QNRPSLAE--LPQAGPETPGLEGLKELMQLCWSSEPKDRPSFQECLPKTDEVFQMVE--N 292

Qy      302 FLEAVIQLKKTKLQSVSSA 320
             : | :   |  |  : |:
Db      293 NMNAAVSTVKDFLSQLKSS 311

RESULT   7
Q9STU7
ID   Q9STU7      PRELIMINARY;      PRT;   963 AA.
AC   Q9STU7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   HYPOTHETICAL 108.0 KDA PROTEIN.
GN   T22A6.310 OR AT4G24480.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Zimmermann W., Grueneisen A., Wambutt R., Bancroft I.,
RA   Mewes H.W., Mayer K.F.X., Lemcke K., Schueller C.;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Zimmermann W., Grueneisen A., Wambutt R., Kalicki J., Wohldmann P.,
RA   Smith A., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL078637; CAB45083.1; -.
DR   EMBL; AL161561; CAB79358.1; -.
DR   HSSP; P06239; 3LCK.
DR   INTERPRO; IPR000719; -.
DR   INTERPRO; IPR001245; -.
DR   INTERPRO; IPR002290; -.
DR   PFAM; PF00069; pkinase; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   963 AA;  107993 MW;  10DD8910F44C140E CRC64;

  Query Match             12.7%;  Score 359.5;  DB 10;  Length 963;
  Best Local Similarity   31.7%;  Pred. No. 1.5e-20;
  Matches   93;  Conservative   60;  Mismatches  105;  Indels   35;  Gaps   11;

Qy       20 DLRYLSRGASGTVSSARHADWR-VQVAVKHLHIHTPLLDSERKDVLRE-----AEILHKA 73
```

```
            :|    |  :|:  :   |:|    |||| | |     | : :: |||      |:  :
Db      668 ELHIKERVGAGSFGTVHRAEWHGSDVAVKILSIQ-DFHDDQFREFLREVCKQAVAIMKRV 726

Qy       74 RFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKT--EYPDVAWPLRFRI-----LH 126
            |   :   :|   |   | |:|||:| |||  |:||    |  |    ||  :     :
Db      727 RHPNVVLFMGAVTERPRLSIITEYLPRGSLFRLIHRPASGELLDQRRRLRMALDVVCAIP 786

Qy      127 EIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPE 186
              | |:||||: ||::| |||: |:|:|  : ||: |||||:::  :   |:|    ||
Db      787 HYAKGLNYLHCLNPPVVHWDLKSPNLLVDKNWTVKVCDFGLSRFKANTFIPSKSVAGTPE 846

Qy      187 GGTIIYMPPE--NYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVS- 243
                 :| ||    ||  :     | |:||: |: ||::: :||:  : :| |:: :|:
Db      847 -----WMAPEFLRGEPTNE-----KSDVYSFGVVLWELITLQQPWNGL-SPAQVVGAVAF 895

Qy      244 QGHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRT 296
            |  | :|   : |        ::||:|: ||  | :||:|   :  |: :|::
Db      896 QNRRLIIPPNTSPV-------LVSLMEACWADEPSQRPAFGSIVDTLKKLLKS 941
```

```
RESULT   8
Q9M8C1
ID   Q9M8C1      PRELIMINARY;      PRT;   773 AA.
AC   Q9M8C1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PROTEIN KINASE, PUTATIVE.
GN   F5E6.5.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA   Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA   Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT   "Arabidopsis thaliana chromosome III BAC F5E6 genomic sequence.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC020580; AAF63631.1; -.
KW   Kinase.
SQ   SEQUENCE   773 AA;  86054 MW;  C1B70D3E6848A340 CRC64;

  Query Match             12.5%;  Score 353;  DB 10;  Length 773;
  Best Local Similarity   32.0%;  Pred. No. 3.8e-20;
  Matches   94;  Conservative   62;  Mismatches  102;  Indels   36;  Gaps   11;

Qy       24 LSRGASGTVSSARHADW-RVQVAVK---HLHIHTPLLDSERKDVLREAEILHKARFSYIF 79
            : :|: |||    |  |    ||||          :::| :::||    :: : |   :
Db      500 VGQGSCGTV---YHGLWFGSDVAVKVFSKQEYSAEVIESFKQEVL----LMKRLRHPNVL 552

Qy       80 PILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMT 139
              :|    |: | ||:|::| |||  || : |   |  |  |  :  :|| |:||||: :
Db      553 LFMGAVTSPQRLCIVSEFLPRGSLFRLLQKSTSKLD--WRRRIHMALDIARGMNYLHHCS 610

Qy      140 PPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPENYE 199
            ||::| |||: |:|:|  : ||:||||||: :  :   |:| |  |:     :| ||
Db      611 PPIIHRDLKSSNLLVDKNWTVKVADFGLSRIKHETYLTSKSGKGTPQ-----WMAPEVL- 664

Qy      200 PGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDI 259
              :   |  | ||||: |: ||: : | |:| : | :|:: :|    : :     :| ||
Db      665 --RNESADEKSDIYSFGVVLWELATEKIPWETL-NSMQVIGAVGFMDQRL----EIPKDI 717

Qy      260 PHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEITFLEAVIQLKKTK 313
                | |||:|| |  :   ||:| : : :|  : | :        :|| : |:
Db      718 D--PRWISLMESCWHSDTKLRPTFQELMDKLRDLQRKY--------MIQFQATR 761

RESULT   9
```

```
Q9S7D5
ID   Q9S7D5      PRELIMINARY;      PRT;   370 AA.
AC   Q9S7D5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PROTEIN KINASE ATN1-LIKE PROTEIN.
GN   T3A5.110 OR F18B3.10.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bloecker H., Mewes H.W., Lemcke K., Mayer K.F.X., Quetier F.,
RA   Salanoubat M.;
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Rieger M., Gabel C., Mueller-Auer S., Schaefer M., Zipp M.,
RA   Mewes H.W., Mayer K.F.X., Lemcke K., Schueller C., Quetier F.,
RA   Salanoubat M.;
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL132979; CAB62442.1; -.
DR   EMBL; AL049862; CAB42902.1; -.
DR   HSSP; P08631; 1AD5.
DR   INTERPRO; IPR000719; -.
DR   INTERPRO; IPR001245; -.
DR   PFAM; PF00069; pkinase; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Kinase.
SQ   SEQUENCE   370 AA;  42328 MW;  A508F716B432804B CRC64;

  Query Match             12.3%;  Score 348.5;  DB 10;  Length 370;
  Best Local Similarity   29.2%;  Pred. No. 3.2e-20;
  Matches   94;  Conservative   63;  Mismatches  106;  Indels   59;  Gaps   13;

Qy       44 VAVKHLHIHT--PLLDSERKDVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNG 101
            |||| :   |   :  : :|   :|  :| | :   |   :| | ||: : |||| :  |
Db       61 VAVKIMDPSTTSAVTKAHKKTFQKEVLLLSKMKHDNIVKFVGACIEPQLI-IVTELVEGG 119

Qy      102 SLNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEF-HV 160
            :|   :| :    |:   | | :  :|:  : ::|  :  ::| ||  :|:|:  :  ||
Db      120 TLQRFMHSRPGPLDLKMSLSFAL--DISRAMEFVH--SNGIIHRDLNPRNLLVTGDLKHV 175

Qy      161 KIADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPENYEP-----GQKSRASIKHDIYSY 215
            |:||||:::       ::|   :  | ||  :| || | |     |:|     | ||||:
Db      176 KLADFGIAR------EETRGGMTC-EAGTSKWMAPEVYSPEPLRVGEKKEYDHKADIYSF 228

Qy      216 AVITWEVLSRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIPHRARMISLIESGWAQ 275
            |:: |:::: ::|| || | | : | |||| ||:: :     |:      : ::|| |||
Db      229 AIVLWQLVTNEEPFPDVPNSLFVPYLVSQGRRPILTKTP---DV-----FVPIVESCWAQ 280

Qy      276 NPDERPSFLKCLIELEPVLRTFEEITFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLNI 335
            :|| || | :  : |  :||                 :: | ||           :   :
Db      281 DPDARPEFKEISVMLTNLLR-----------------RMSSDSS-----------IGTTL 312

Qy      336 P---VNHGPQEESCGSSQLHEN 354
            |      |  |||  |  | |:
Db      313 PDGEAYEGEMEESENSPLLQEH 334
```

```
RESULT  10
Q9M8C3
ID   Q9M8C3      PRELIMINARY;      PRT;   763 AA.
AC   Q9M8C3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PROTEIN KINASE, PUTATIVE.
GN   F5E6.3.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA   Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA   Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT   "Arabidopsis thaliana chromosome III BAC F5E6 genomic sequence.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC020580; AAF63629.1; -.
KW   Kinase.
SQ   SEQUENCE   763 AA;  85593 MW;  F6787110603C597C CRC64;

  Query Match             12.3%;  Score 348.5;  DB 10;  Length 763;
  Best Local Similarity   31.9%;  Pred. No. 8.7e-20;
  Matches   99;  Conservative   57;  Mismatches  109;  Indels   45;  Gaps   13;

Qy     24 LSRGASGTVSSARHADW-RVQVAVKHL---HIHTPLLDSERKDVLREAEILHKARFSYIF 79
          : :|: |||    |  |    |||| :        :: | |::|     :: : |   :
Db    452 IGQGSCGTV---YHGLWFGSDVAVKLISKQEYSEEVIQSFRQEV----SLMQRLRHPNVL 504

Qy     80 PILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMT 139
            :|    |: | ||:|::| |||  || |     |  |  |  :  :|| |:||||  :
Db    505 LFMGAVTLPQGLCIVSEFLPRGSLFRLLQRNMSKLD--WRRRINMALDIARGMNYLHRCS 562

Qy    140 PPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPENYE 199
          ||::| |||: |:|:|    ||:||||||: :  :   |:| |  |:     :| ||
Db    563 PPIIHRDLKSSNLLVDKNLTVKVADFGLSRIKHHTYLTSKSGKGMPQ-----WMAPEVL- 616

Qy    200 PGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVSQGHRPVINEE-SLPYD 258
            :   |  | ||||: |: ||: : | |:|:: | :|:: :|       :|:   :| |
Db    617 --RNESADEKSDIYSFGVVLWELATEKIPWENL-NSMQVIGAVG-----FMNQRLEIPKD 668

Qy    259 IPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEITFLEAVIQLKKTKLQSVS 318
          |      |||||| | ::   ||:| : :  |  : | :         || : |:
Db    669 ID--PDWISLIESCWHRDAKLRPTFQELMERLRDLQRKY--------TIQFQATRAA--- 715

Qy    319 SAIHLCDKKK 328
              | ||  |
Db    716 ----LSDKNK 721

RESULT  11
O65833
ID   O65833      PRELIMINARY;      PRT;   981 AA.
AC   O65833;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   TCTR2 PROTEIN.
GN   TCTR2.
OS   Lycopersicon esculentum (Tomato).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC   Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4081;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. AILSA CRAIG;
RA   Lin Z., Hackett R.M., Payton S., Grierson D.;
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ005077; CAA06334.1; -.
DR   HSSP; P12931; 1FMK.
DR   MENDEL; 29910; Lyces;2342;29910.
DR   INTERPRO; IPR000568; -.
DR   INTERPRO; IPR000719; -.
DR   INTERPRO; IPR002290; -.
DR   PFAM; PF00069; pkinase; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS00449; ATPASE_A; UNKNOWN_1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Kinase.
SQ   SEQUENCE   981 AA;  107175 MW;  04522B40F3425068 CRC64;

  Query Match             12.0%;  Score 340;  DB 10;  Length 981;
  Best Local Similarity   30.8%;  Pred. No. 6e-19;
  Matches   90;  Conservative   56;  Mismatches  108;  Indels   38;  Gaps   11;

Qy     13 IPYHKLADLRYLSRGASGTVSSARHADWR-VQVAVKHLHIHTPLLDSE-----RKDVLRE 66
          ||:  |     :  |: | |    ||||  :||||         || :       :  ||
Db    695 IPWEDLVIGERIGLGSYGEV---YHADWNGTEVAVK------KFLDQDFSGAALAEFKRE 745

Qy     67 AEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILH 126
            |: : |   :   :|    |  | |:||::| |||  ::||     |    ::  :
Db    746 VRIMRRLRHPNVVRFMGAITRPPHLSIITEFLPRGSLYRIIHRPHFQIDERQKIKMAL-- 803

Qy    127 EIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPE 186
          ::| |:: ||   | ::| |||: |:|:| :::||: |||||: :  :   |:|:   ||
Db    804 DVAKGMDCLHTSNPTIVHRDLKSPNLLVDTDWNVKVCDFGLSRLKHNTFLSSKSTAGTPE 863

Qy    187 GGTIIYMPPE--NYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVS- 243
               :| ||    ||  :     | ||||: || ||: : : |:  : ||:|:: :|
Db    864 -----WMAPEVLRNEPSNE-----KCDIYSFGVILWELATLRLPWSGM-NPMQVVGAVGF 912

Qy    244 QGHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLR 295
          |  |  | :|  |        :  :|   |  :|: |||| :  : | |: |
Db    913 QNKRLEIPKELDPI-------VARIIWECWQTDPNLRPSFAQLTVALTPLQR 957

RESULT  12
Q9SFN9
ID   Q9SFN9      PRELIMINARY;      PRT;  1030 AA.
AC   Q9SFN9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   HYPOTHETICAL 112.2 KDA PROTEIN.
GN   F25P22.8.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA   Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA   Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT   "Arabidopsis thaliana chromosome I BAC F25P22 genomic sequence.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC012679; AAF24836.1; -.
DR   HSSP; P00523; 2PTK.
DR   INTERPRO; IPR000104; -.
DR   INTERPRO; IPR000719; -.
DR   INTERPRO; IPR001245; -.
DR   INTERPRO; IPR002290; -.
DR   PFAM; PF00069; pkinase; 1.
DR   PRINTS; PR00109; TYRKINASE.
```

```
DR   PRINTS; PR00308; ANTIFREEZEI.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   1030 AA;  112205 MW;  9BD7125FB1DF1B38 CRC64;

  Query Match             11.9%;  Score 337.5;  DB 10;  Length 1030;
  Best Local Similarity   34.8%;  Pred. No. 1e-18;
  Matches   93;  Conservative   39;  Mismatches  108;  Indels   27;  Gaps   10;

Qy         30 GTVSSARHADWR-VQVAVKHLHIHTPLLDSERKDVLREAEILHKARFSYIFPILGICNEP 88
              |:       ||   :||||   :   |     ::   |  |: | |   |   :|    |
Db        757 GSYGEVYRGDWHGTEVAVKKF-LDQDLTGEALEEFRSEVRIMKKLRHPNIVLFMGAVTRP 815

Qy         89 EFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLK 148
                | ||||::| |||  |:||     |     | |:  : | |:||||:  | ::| |||
Db        816 PNLSIVTEFLPRGSLYRLIHRPNNQLDER--RRLRMALDAARGMNYLHSCNPMIVHRDLK 873

Qy        149 TQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPE--NYEPGQKSRA 206
              : |:|:|  : ||: |||||:   |  |   ||||    ||  :| ||    ||     |
Db        874 SPNLLVDKNWVVKVCDFGLSR---MKHSTYLSSKST--AGTAEWMAPEVLRNEP-----A 923

Qy        207 SIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIPH--RAR 264
                | |:||| || ||: : :||:  : ||:|:: :|   || :        |||
Db        924 DEKCDVYSYGVILWELFTLQQPWGKM-NPMQVVGAVGFQHRRL--------DIPDFVDPA 974

Qy        265 MISLIESGWAQNPDERPSFLKCLIELE 291
              :  ||   |  :   |||| : :  |:
Db        975 IADLISKCWQTDSKLRPSFAEIMASLK 1001

RESULT  13
O82754
ID   O82754      PRELIMINARY;      PRT;   736 AA.
AC   O82754;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PUTATIVE SERINE/THREONINE KINASE.
GN   F7H19.240 OR AT4G23050.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Peters S.A., van Staveren M., Dirkse W., Stiekema W., Bancroft I.,
RA   Mewes H.W., Mayer K.F.X., Schueller C., Bevan M.;
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Peters S.A., van Staveren M., Dirkse W., Stiekema W., Mewes H.W.,
RA   Lemcke K., Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL031018; CAA19821.1; -.
DR   EMBL; AL161558; CAB79260.1; -.
DR   HSSP; P00523; 2PTK.
DR   INTERPRO; IPR000014; -.
DR   INTERPRO; IPR000719; -.
DR   INTERPRO; IPR002290; -.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00989; PAS; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Kinase.
```

```
SQ   SEQUENCE   736 AA;  82281 MW;  E47D333311805B52 CRC64;

  Query Match             11.9%;  Score 335.5;  DB 10;  Length 736;
  Best Local Similarity   32.9%;  Pred. No. 9.3e-19;
  Matches   79;  Conservative   50;  Mismatches   76;  Indels   35;  Gaps   10;

Qy         65 REAEILHKARFSYIFPILG-ICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPL--- 120
              :|  |: | |   :   :| :| | |   |: |||| ||| ::||   :      ||
Db        512 KEINIMKKLRHPNVLLFMGAVCTE-EKSAIIMEYMPRGSLFKILHNTNQ------PLDKK 564

Qy        121 -RFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSR 179
               | |:  ::| |:||||   ||::| |||: |:|:|  ::||: |||||||:  :  ::
Db        565 RRLRMALDVARGMNYLHRRNPPIVHRDLKSSNLLVDKNWNVKVGDFGLSKWKNATFLSTK 624

Qy        180 SSKSAPEGGTIIYMPPE--NYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQ 237
              | |  |:     :| ||    ||  :     | |::|: || ||:::   |: |  | :|
Db        625 SGKGTPQ-----WMAPEVLRSEPSNE-----KCDVFSFGVILWELMTTLVPW-DRLNSIQ 673

Qy        238 IMYSVSQGHRPVINEESLPYDIPH--RARMISLIESGWAQNPDERPSFLKCLIELEPVLR 295
              ::  |    | :        |:|     |: |:|:  |  :| :|||| : : ::  : |
Db        674 VVGVVGFMDRRL--------DLPEGLNPRIASIIQDCWQTDPAKRPSFEELISQMMSLFR 725

RESULT  14
Q9Y2V6
ID   Q9Y2V6      PRELIMINARY;      PRT;   835 AA.
AC   Q9Y2V6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   HYPOTHETICAL 92.9 KDA PROTEIN.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=HEART;
RA   Wei Y.J., Ding J.F., Xiong H., Zhou Y., Hui R.T., Liew C.C.;
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF116826; AAD29632.1; -.
DR   HSSP; P00523; 2PTK.
DR   INTERPRO; IPR000719; -.
DR   INTERPRO; IPR001245; -.
DR   INTERPRO; IPR002110; -.
DR   PFAM; PF00023; ank; 10.
DR   PFAM; PF00069; pkinase; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   835 AA;  92850 MW;  3B21484B434F46E8 CRC64;

  Query Match             11.8%;  Score 334.5;  DB 4;  Length 835;
  Best Local Similarity   28.1%;  Pred. No. 1.3e-18;
  Matches  116;  Conservative   76;  Mismatches  170;  Indels   51;  Gaps   19;

Qy          8 SALPTIPYHKLADLRY---LSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSERKDVL 64
              : ||:  : :|::: :   :  |: | |   |  :    ||:|    :|    |:
Db        450 AGLPSHFHLQLSEIEFHEIIGSGSFGKVYKGRCRN--KIVAIKRYRANTYCSKSDVDMFC 507

Qy         65 REAEILHKARFSYIFPILGIC-NEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFR 123
              ||  || :     :   :| | |:|    |||:|:  |||  ||| :    |:   |
Db        508 REVSILCQLNHPCVIQFVGACLNDPSQFAIVTQYISGGSLFSLLHEQKRILDLQSKLIIA 567

Qy        124 ILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKS 183
              :  ::| |: ||||:| |::| || : ||||  : |  :|||| |:: : || :   :|
Db        568 V--DVAKGMEYLHNLTQPIIHRDLNSHNILLYEDGHAVVADFGESRF-LQSLDEDNMTK- 623

Qy        184 APEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVS 243
                : | : :| || :   | :| :|| |::|||:  ||:|: : ||  :  |     ::
```

```
Db    624 --QPGNLRWMAPEVFT--QCTRYTIKADVFSYALCLWEILTGEIPFAHL-KPAAAAADMA 678

Qy    244 QGH-RPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEITF 302
            | || |      | ||    : ||:  ||   |: || | : :::||  |   |
Db    679 YHHIRPPIG-----YSIP--KPISSLLIRGWNACPEGRPEFSEVVMKLEECLCNIE---L 728

Qy    303 LEAVIQLKKTKLQSVSSAIHLCDKK--------------KMELSLNIPVNHGPQEESCG- 347
          :          |   ||:  | ::               ::| :||   ::    :| |
Db    729 MSPASSNSSGSLSPSSSSDCLVNRGGPGRSHVAALRSRFELEYALNAR-SYAALSQSAGQ 787

Qy    348 -SSQLHENSGSPETSRSLP-APQDN-DFLSRKAQDCYFMKLHHCPGNHSWDST 397
           |||        |  |||   | |   ::|      :|   | |  : |:: :
Db    788 YSSQ---GLSLEEMKRSLQYTPIDKYGYVSDPMSSMHF---HSCRNSSSFEDS 834

RESULT  15
Q9LVQ9
ID   Q9LVQ9      PRELIMINARY;      PRT;   405 AA.
AC   Q9LVQ9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PROTEIN KINASE ATN1-LIKE PROTEIN.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=COLUMBIA;
RA   Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA   Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT   features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT   clones.";
RL   DNA Res. 7:31-63(2000).
DR   EMBL; AB018119; BAA97277.1; -.
KW   Kinase.
SQ   SEQUENCE   405 AA;  46012 MW;  C7C4CF29E8DA3111 CRC64;


  Query Match            11.8%;  Score 332.5;  DB 10;  Length 405;
  Best Local Similarity  30.1%;  Pred. No. 7.1e-19;
  Matches  112;  Conservative   64;  Mismatches  135;  Indels   61;  Gaps   18;

Qy     12 TIPYHKLADLRYLSRG---ASGTVSSARHADWR--VQVAVKHLHIHTP-----LLDSERK 61
          ||    | |:: :| |     |: |:     :|  |  |:||   |  |     |   :||
Db     59 TINTELLVDVKDISIGDFIGEGSSSTVYRGLFRRVVPVSVK---IFQPKRTSALSIEQRK 115

Qy     62 DVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYP-DVAWPL 120
             ||  :| | |   |   :| | ||: : |:|| |   :| : :      | |:   :
Db    116 KFQREVLLLSKFRHENIVRFIGACIEPKLM-IITELMEGNTLQKFMLSVRPKPLDLKLSI 174

Qy    121 RFRILHEIALGVNYLHNMTPPLLHHDLKTQNILL-DNEFHVKIADFGLSKWRMMSLSQSR 179
           | :  :|| |: :|:     ::| |||  |:||  :: |||:|||||::       :
Db    175 SFAL--DIARGMEFLN--ANGIIHRDLKPSNMLLTGDQKHVKLADFGLAR-------EET 223

Qy    180 SSKSAPEGGTIIYMPPENY-----EPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTN 234
                |  ||  :| || :     | |:|     |  |:||:|:: ||:|: | ||:   |
Db    224 KGFMTFEAGTYRWMAPELFSYDTLEIGEKKHYDHKVDVYSFAIVFWELLTNKTPFKGKNN 283

Qy    235 PLQIMYSVSQGHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVL 294
           : : |: |:  || :  |:||  :      :|:::| ||:||| || | :    |  :|
Db    284 -IFVAYAASKNQRPSV--ENLPEGV------VSILQSCWAENPDARPEFKEITYSLTNLL 334

Qy    295 RTFEEITFLEAVIQLKKTKL---QSVSSAIH---LCD------------KKKMELSLNIP 336
          |:    |  :|     |  :    | || :    :||            |||    :|:
Db    335 RSLSSDT--DATSSNSKANIATEDSTSSLVQERVVCDCPGLKMSKTKKLKKKTNKLMNMI 392

Qy    337 VNHGPQEESCGS 348
          |      :|| |
Db    393 VPFLKIFKSCMS 404
```

```
Search completed: June 14, 2001, 08:06:41
Job time: 132 sec
```

```
                     GenCore version 4.5
              Copyright (c) 1993 - 2000  Compugen Ltd.


OM protein - protein search, using sw model

Run on:          June 14, 2001, 08:04:28 ; Search time 51.9 Seconds
                                            (without alignments)
                                            594.762 Million cell updates/sec

Title:           US-09-445-223-1
Perfect score:   2829
Sequence:        1 MNGEAICSALPTIPYHKLAD..........PEILVVSRSPSLNLLQNKSM 540

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters:      390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_0401:*
              1:  /SIDS6/gcgdata/geneseq/geneseqp/AA1980.DAT:*
              2:  /SIDS6/gcgdata/geneseq/geneseqp/AA1981.DAT:*
              3:  /SIDS6/gcgdata/geneseq/geneseqp/AA1982.DAT:*
              4:  /SIDS6/gcgdata/geneseq/geneseqp/AA1983.DAT:*
              5:  /SIDS6/gcgdata/geneseq/geneseqp/AA1984.DAT:*
              6:  /SIDS6/gcgdata/geneseq/geneseqp/AA1985.DAT:*
              7:  /SIDS6/gcgdata/geneseq/geneseqp/AA1986.DAT:*
              8:  /SIDS6/gcgdata/geneseq/geneseqp/AA1987.DAT:*
              9:  /SIDS6/gcgdata/geneseq/geneseqp/AA1988.DAT:*
             10:  /SIDS6/gcgdata/geneseq/geneseqp/AA1989.DAT:*
             11:  /SIDS6/gcgdata/geneseq/geneseqp/AA1990.DAT:*
             12:  /SIDS6/gcgdata/geneseq/geneseqp/AA1991.DAT:*
             13:  /SIDS6/gcgdata/geneseq/geneseqp/AA1992.DAT:*
             14:  /SIDS6/gcgdata/geneseq/geneseqp/AA1993.DAT:*
             15:  /SIDS6/gcgdata/geneseq/geneseqp/AA1994.DAT:*
             16:  /SIDS6/gcgdata/geneseq/geneseqp/AA1995.DAT:*
             17:  /SIDS6/gcgdata/geneseq/geneseqp/AA1996.DAT:*
             18:  /SIDS6/gcgdata/geneseq/geneseqp/AA1997.DAT:*
             19:  /SIDS6/gcgdata/geneseq/geneseqp/AA1998.DAT:*
             20:  /SIDS6/gcgdata/geneseq/geneseqp/AA1999.DAT:*
             21:  /SIDS6/gcgdata/geneseq/geneseqp/AA2000.DAT:*
             22:  /SIDS6/gcgdata/geneseq/geneseqp/AA2001.DAT:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                              SUMMARIES

                  %
Result          Query
  No.   Score   Match Length DB  ID                     Description
 ----------------------------------------------------------------------
    1    2829   100.0    540  20  W92795                 Human B1 protein.
    2    2823    99.8    540  21  Y68774                 Amino acid sequenc
    3    2823    99.8    544  21  B43570                 Human cancer assoc
    4    2817    99.6    540  20  Y31140                 Human CARD-3 prote
    5    2774    98.1    531  21  Y59404                 Human RICK protein
    6    2497    88.3    478  21  Y59405                 Human RICK protein
    7    1475    52.1    284  21  Y59406                 Human RICK protein
    8     867    30.6    167  21  Y59407                 Human RICK protein
    9     797    28.2    153  21  B58938                 Breast and ovarian
   10     577    20.4    786  21  Y69163                 Amino acid sequenc
   11     576    20.4    787  21  Y76079                 Murine protein kin
```

10/504 2%

```
   12     576    20.4    787  22  B56018                 Skin cell protein,
   13     574    20.3    590  21  Y76123                 Murine RIP protein
   14     574    20.3    590  22  B56062                 Skin cell protein,
   15     574    20.3    763  21  Y79154                 Mouse protein kina
   16     536    18.9    536  21  Y76007                 Murine RIP protein
   17     536    18.9    536  22  B55946                 Skin cell protein,
   18     394    13.9    656  18  W04627                 Mouse receptor int
   19     389    13.8    518  21  B01530                 Kinase of death (K
   20     388    13.7    518  21  B01524                 Kinase of death (K
   21     388    13.7    656  20  W80994                 Human receptor int
   22   386.5    13.7    420  21  B01529                 Kinase of death (K
   23     385    13.6    518  21  B01526                 Kinase of death (K
   24     380    13.4    518  21  B01525                 Kinase of death (K
   25     376    13.3    821  16  R80574                 Arabidopsis CTR1 p
   26     376    13.3    821  18  W17938                 Constitutive tripl
   27     376    13.3    821  22  B50438                 Arabidopsis thalia
   28     375    13.3    436  21  Y45043                 Human Apop3 (1-436
   29     375    13.3    519  21  Y45042                 Human Apop3 protei
   30     372    13.1    261  21  B01527                 Kinase of death (K
   31     372    13.1    821  15  R46723                 Arabidopsis thalia
   32   371.5    13.1    671  18  W04628                 Human receptor int
   33     370    13.1    671  18  W15461                 Human receptor int
   34     369    13.0    519  21  Y45046                 Human Apop3 (K50D)
   35     369    13.0    671  21  Y78502                 Human RIP-1 amino
   36     359    12.7    933  22  B50437                 Arabidopsis thalia
   37     347    12.3    485  21  B18658                 A human regulator
   38   341.5    12.1    970  22  B50443                 Barley EDR1.  Hord
   39     340    12.0    982  22  B50439                 Tomato TCTR2.  Lyc
   40   334.5    11.8    835  21  B01470                 Human CARK (Cardia
   41   334.5    11.8    835  22  B65674                 Novel protein kina
   42     331    11.7    437  21  Y45047                 Human Apop3 (82-51
   43   330.5    11.7    903  22  B50440                 Rice EDR1.  Oryza
   44   322.5    11.4    251  21  Y45044                 Human Apop3 (1-251
   45     316    11.2    984  21  B25544                 Eucalyptus grandis

                              ALIGNMENTS


RESULT   1
W92795
ID   W92795 standard; Protein; 540 AA.
XX
AC   W92795;
XX
DT   07-MAY-1999  (first entry)
XX
DE   Human B1 protein.
XX
KW   B1 protein; intracellular mediator; modulator; inflammation; cell death;
KW   cell survival pathway; intracellular signalling; AIDS; cancer; human.
XX
OS   Homo sapiens.
XX
PN   WO9855507-A2.
XX
PD   10-DEC-1998.
XX
PF   01-JUN-1998;   98WO-IL00255.
XX
PR   11-SEP-1997;   97IL-0121746.
PR   05-JUN-1997;   97IL-0121011.
PR   30-JUN-1997;   97IL-0121199.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Boldin M,  Malinin N,  Wallach D;
XX
DR   WPI; 1999-070258/06.
DR   N-PSDB; X02558.
XX
PT   New B1 protein regulates cell death and cell survival pathways -
PT   derivatives, DNA and antibodies, also regulate intracellular
```

```
PT   inflammation; for treating AIDS, cancer
XX
PS   Claim 4; Fig 3A; 90pp; English.
XX
CC   This invention describes the isolation of a novel human B1 protein which
CC   can interact with, intracellular mediators or modulators of inflammation,
CC   cell death and/or cell survival pathways, directly or indirectly. Cells
CC   can be modulated or mediated in inflammation, cell death or cell survival
CC   pathways or another intracellular signalling activity using B1.
CC   Conditions such as AIDS and cancer can be treated using B1. Antibodies,
CC   oligonucleotides and ribozymes can also be used to regulate the above
CC   pathways.
XX
SQ   Sequence   540 AA;

  Query Match             100.0%;  Score 2829;  DB 20;  Length 540;
  Best Local Similarity   100.0%;  Pred. No. 5.2e-257;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 MNGEAICSALPTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 mngeaicsalptipyhkladlrylsrgasgtvssarhadwrvqvavkhlhihtplldser 60

Qy       61 KDVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 kdvlreaeilhkarfsyifpilgicnepeflgivteympngslnellhrkteypdvawpl 120

Qy      121 RFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRS 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 rfrilheialgvnylhnmtppllhhdlktqnilldnefhvkiadfglskwrmmslsqsrs 180

Qy      181 SKSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMY 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 sksapeggtiiymppenyepgqksrasikhdiysyavitwevlsrkqpfedvtnplqimy 240

Qy      241 SVSQGHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEI 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 svsqghrpvineeslpydiphrarmisliesgwaqnpderpsflkclielepvlrtfeei 300

Qy      301 TFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLNIPVNHGPQEESCGSSQLHENSGSPET 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 tfleaviqlkktklqsvssaihlcdkkkmelslnipvnhgpqeescgssqlhensgspet 360

Qy      361 SRSLPAPQDNDFLSRKAQDCYFMKLHHCPGNHSWDSTISGSQRAAFCDHKTTPCSSAIIN 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 srslpapqdndflsrkaqdcyfmklhhcpgnhswdstisgsqraafcdhkttpcssaiin 420

Qy      421 PLSTAGNSERLQPGIAQQWIQSKREDIVNQMTEACLNQSLDALLSRDLIMKEDYELVSTK 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 plstagnserlqpgiaqqwiqskredivnqmteaclnqsldallsrdlimkedyelvstk 480

Qy      481 PTRTSKVRQLLDTTDIQGEEFAKVIVQKLKDNKQMGLQPYPEILVVSRSPSLNLLQNKSM 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 ptrtskvrqlldttdiqgeefakvivqklkdnkqmglqpypeilvvsrspslnllqnksm 540


RESULT    2
Y68774
ID   Y68774 standard; Protein; 540 AA.
XX
AC   Y68774;
XX
DT   16-MAY-2000  (first entry)
XX
DE   Amino acid sequence of a human phosphorylation effector PHSP-6.
XX
KW   Human; phosphorylation effector; PHSP; proliferative disorder;
KW   immune disorder; neuronal disorder.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          18..287
FT                   /note= "protein kinase family signature sequence"
FT   Modified-site   23
FT                   /note= "potential phosphorylation site"
FT   Modified-site   34
FT                   /note= "potential phosphorylation site"
FT   Modified-site   58
FT                   /note= "potential phosphorylation site"
FT   Modified-site   100
FT                   /note= "potential glycosylation site"
FT   Modified-site   102
FT                   /note= "potential phosphorylation site"
FT   Modified-site   180
FT                   /note= "potential phosphorylation site"
FT   Modified-site   183
FT                   /note= "potential phosphorylation site"
FT   Modified-site   207
FT                   /note= "potential phosphorylation site"
FT   Modified-site   224
FT                   /note= "potential phosphorylation site"
FT   Modified-site   267
FT                   /note= "potential phosphorylation site"
FT   Modified-site   296
FT                   /note= "potential phosphorylation site"
FT   Modified-site   301
FT                   /note= "potential phosphorylation site"
FT   Modified-site   360
FT                   /note= "potential phosphorylation site"
FT   Modified-site   374
FT                   /note= "potential phosphorylation site"
FT   Modified-site   391
FT                   /note= "potential glycosylation site"
FT   Modified-site   401
FT                   /note= "potential phosphorylation site"
FT   Modified-site   428
FT                   /note= "potential phosphorylation site"
FT   Modified-site   442
FT                   /note= "potential phosphorylation site"
FT   Modified-site   457
FT                   /note= "potential glycosylation site"
FT   Modified-site   478
FT                   /note= "potential phosphorylation site"
FT   Modified-site   478
FT                   /note= "potential phosphorylation site"
FT   Modified-site   484
FT                   /note= "potential phosphorylation site"
FT   Modified-site   537
FT                   /note= "potential glycosylation site"
XX
PN   WO200006728-A2.
XX
PD   10-FEB-2000.
XX
PF   28-JUL-1999;   99WO-US17132.
XX
PR   28-JUL-1998;   98US-0123494.
PR   14-SEP-1998;   98US-0152814.
PR   14-OCT-1998;   98US-0173482.
PR   03-NOV-1998;   98US-0106889.
PR   19-NOV-1998;   98US-0109093.
PR   22-DEC-1998;   98US-0113796.
PR   12-JAN-1999;   99US-0173482.
PR   12-JAN-1999;   99US-0229005.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Hillman JL,  Lal P,  Tang YT,  Corley NC,  Guegler KJ,  Baughn MR;
PI   Patterson C,  Bandman O,  Au-Young J,  Gorgone GA,  Yue H,  Azimzai Y;
PI   Reddy R,  Lu DAM,  Shih LL;
XX
DR   WPI; 2000-183125/16.
DR   N-PSDB; Z46143.
XX
```

```
PT   New human phosphorylation effectors useful for the diagnosis, treatment
PT   and prevention of proliferative, immune and neuronal disorders -
XX
PS   Claim 1; Page 84-85; 142pp; English.
XX
CC   Y68769-95 and Y68797-99 represent human phosphorylation effectors (PHSP),
CC   designated PHSP1-PHSP31 (the protein sequence for PHSP28 is not given
CC   in the specification). The sequences were isolated from cDNA libraries
CC   prepared from various human tissues. The PHSP proteins are useful for
CC   the diagnosis, treatment and prevention of proliferative disorders,
CC   immune disorders and neuronal disorders. The PHSP proteins form
CC   pharmaceutical compositions which useful for treating or preventing
CC   disorders associated with decreased PHSP expression/activity. PHSP
CC   antagonists are useful for treating or preventing disorders associated
CC   with increased PHSP expression/activity.
XX
SQ   Sequence   540 AA;

  Query Match             99.8%;  Score 2823;  DB 21;  Length 540;
  Best Local Similarity   99.8%;  Pred. No. 1.9e-256;
  Matches  539;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1 MNGEAICSALPTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 mngeaicsalptipyhkladlrylsrgasgtvssarhadwrvqvavkhlhihtplldser 60

Qy       61 KDVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPL 120
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db       61 kdvlreaeilhkarfsyilpilgicnepeflgivteympngslnellhrkteypdvawpl 120

Qy      121 RFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRS 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 rfrilheialgvnylhnmtppllhhdlktqnilldnefhvkiadfglskwrmmslsqsrs 180

Qy      181 SKSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMY 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 sksapeggtiiympppenyepgqksrasikhdiysyavitwevlsrkqpfedvtnplqimy 240

Qy      241 SVSQGHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEI 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 svsqghrpvineeslpydiphrarmisliesgwaqnpderpsflkclielepvlrtfeei 300

Qy      301 TFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLNIPVNHGPQEESCGSSQLHENSGSPET 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 tfleaviqlkktklqsvssaihlcdkkkmelslnipvnhgpqeescgssqlhensgspet 360

Qy      361 SRSLPAPQDNDFLSRKAQDCYFMKLHHCPGNHSWDSTISGSQRAAFCDHKTTPCSSAIIN 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 srslpapqdndflsrkaqdcyfmklhhcpgnhswdstisgsqraafcdhkttpcssaiin 420

Qy      421 PLSTAGNSERLQPGIAQQWIQSKREDIVNQMTEACLNQSLDALLSRDLIMKEDYELVSTK 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 plstagnserlqpgiaqqwiqskredivnqmteaclnqsldallsrdlimkedyelvstk 480

Qy      481 PTRTSKVRQLLDTTDIQGEEFAKVIVQKLKDNKQMGLQPYPEILVVSRSPSLNLLQNKSM 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 ptrtskvrqlldttdiqgeefakvivqklkdnkqmglqpypeilvvsrspslnllqnksm 540

RESULT   3
B43570
ID   B43570 standard; Protein; 544 AA.
XX
AC   B43570;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human cancer associated protein sequence SEQ ID NO:1015.
XX
KW   Human; cancer associated gene; cancer antigen; detection; cancer;
KW   diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW   antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW   antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW   dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW   vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW   immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW   allergic reaction; graft versus host disease; organ rejection;
KW   haemostatic; thrombolytic; cardiovascular disorder; infection;
KW   neurological disease; drug screening.
XX
OS   Homo sapiens.
XX
PN   WO200055350-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05882.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587533/55.
DR   N-PSDB; C77779.
XX
PT   Novel isolated nucleic acids comprising sequences encoding peptides
PT   useful for treating or diagnosing e.g. cancer -
XX
PS   Claim 11; Page 1595-1597; 2352pp; English.
XX
CC   C77607 to C78448 encode the human cancer associated proteins given in
CC   B43398 to B44239. The proteins can have activities based on the tissues
CC   and cells the genes are expressed in. Example of activities include:
CC   cytostatic; proliferative; vulnerary; immunomodulator; antidiabetic;
CC   antiasthmatic; antirheumatic; antiarthritic; antiinflammatory;
CC   antithyroid; antiallergic; antibacterial; antiviral; dermatological;
CC   neuroprotective; cardiant; thrombolytic; coagulant; nootropic;
CC   vasotropic; antipsoriatic and antiangiogenic. The polynucleotides and
CC   polypeptides can be used for preventing, treating or ameliorating medical
CC   conditions and diagnosing pathological conditions. Polynucleotides,
CC   polypeptides, antibodies, agonists and antagonists from the present
CC   invention may be used to treat immune disorders by activating or
CC   inhibiting the proliferation, differentiation or mobilisation of immune
CC   cells, to treat disorders of haematopoietic cells, autoimmune disorders,
CC   allergic reactions, graft versus host disease and organ rejection,
CC   modulate haemostatic or thrombolytic activity, modulate inflammation,
CC   cancers, cardiovascular disorders, neurological disease and bacterial or
CC   viral infections. The peptides, nucleotides, antibodies, agonists and
CC   antagonists may be also be used in drug screens. C78449 to C78457 and
CC   B44240 represent sequences used in the exemplification of the present
CC   invention.
XX
SQ   Sequence   544 AA;

  Query Match             99.8%;  Score 2823;  DB 21;  Length 544;
  Best Local Similarity   99.8%;  Pred. No. 1.9e-256;
  Matches  539;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1 MNGEAICSALPTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5 mngeaicsalptipyhkladlrylsrgasgtvssarhadwrvqvavkhlhihtplldser 64

Qy       61 KDVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPL 120
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db       65 kdvlreaeilhkarfsyilpilgicnepeflgivteympngslnellhrkteypdvawpl 124

Qy      121 RFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRS 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      125 rfrilheialgvnylhnmtppllhhdlktqnilldnefhvkiadfglskwrmmslsqsrs 184

Qy      181 SKSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMY 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     185 sksapeggtiiymppenyepgqksrasikhdiysyavitwevlsrkqpfedvtnplqimy 244

Qy     241 SVSQGHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEI 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     245 svsqghrpvineeslpydiphrarmisliesgwaqnpderpsflkclielepvlrtfeei 304

Qy     301 TFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLNIPVNHGPQEESCGSSQLHENSGSPET 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     305 tfleaviqlkktklqsvssaihlcdkkkmelslnipvnhgpqeescgssqlhensgspet 364

Qy     361 SRSLPAPQDNDFLSRKAQDCYFMKLHHCPGNHSWDSTISGSQRAAFCDHKTTPCSSAIIN 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     365 srslpapqdndflsrkaqdcyfmklhhcpgnhswdstisgsqraafcdhkttpcssaiin 424

Qy     421 PLSTAGNSERLQPGIAQQWIQSKREDIVNQMTEACLNQSLDALLSRDLIMKEDYELVSTK 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     425 plstagnserlqpgiaqqwiqskredivnqmteaclnqsldallsrdlimkedyelvstk 484

Qy     481 PTRTSKVRQLLDTTDIQGEEFAKVIVQKLKDNKQMGLQPYPEILVVSRSPSLNLLQNKSM 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     485 ptrtskvrqlldttdiqgeefakvivqklkdnkqmglqpypeilvvsrspslnllqnksm 544


RESULT    4
Y31140
ID   Y31140 standard; Protein; 540 AA.
XX
AC   Y31140;
XX
DT   25-OCT-1999  (first entry)
XX
DE   Human CARD-3 protein.
XX
KW   CARD-3; caspase recruitment domain; CARD-4; regulation; detection;
KW   caspase activation; detection; screening; therapy; diagnosis; disease;
KW   apoptotic cell death; Fas/APO-1 receptor complex; TNF receptor complex;
KW   cancer; follicular lymphoma; carcinoma; p53 mutation; viral infection;
KW   hormone-dependent tumour; autoimmune disorder; Alzheimer's disease;
KW   systemic lupus erythematosis; immune-mediated glomerulonephritis; stroke;
KW   Parkinson's disease; amyotrophic lateral sclerosis; retinitis pigmentosa;
KW   spinal muscular dystrophy; cerebellar degeneration; anaemia; drug;
KW   myelodysplastic syndrome; myocardial infarction; cell proliferation;
KW   cell differentiation; cell survival; CARD-4L; CARD-4S; CARD-4Y;
KW   CARD-4Z; human.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          1..300
FT                   /note= "predicted kinase domain"
FT   Domain          301..431
FT                   /note= "predicted linker domain"
FT   Domain          432..540
FT                   /note= "predicted CARD domain"
XX
PN   WO9940102-A1.
XX
PD   12-AUG-1999.
XX
PF   05-FEB-1999;   99WO-US02544.
XX
PR   08-DEC-1998;   98US-0207359.
PR   06-FEB-1998;   98US-0019942.
PR   17-JUN-1998;   98US-0099041.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Bertin J;
XX
DR   WPI; 1999-494269/41.
DR   N-PSDB; Z09246.
XX
PT   Novel CARD-3 and CARD-4 genes and polypeptides used or treating
PT   regulation of cellular proliferation and differentiation and cell
PT   survival
XX
PS   Example 2; Fig 2; 181pp; English.
XX
CC   This invention describes the isolation of novel human caspase
CC   recruitment domain, CARD-3 and CARD-4 polynucleotides and proteins and a
CC   partial murine CARD-4L protein and genes. The genes and proteins of
CC   the invention are involved in the regulation of caspase activation.
CC   The caspase recruitment domain (CARD) polynucleotides, polypeptides,
CC   homologues and antibodies can be used in screening assays, detection
CC   assays, predictive medicine and therapeutic and prophylactic methods of
CC   treatment. The methods may be used to diagnose and treat patients which
CC   are suffering from a disorder associated with abnormal level or rate of
CC   apoptotic cell death, abnormal activity of the Fas/APO-1 receptor
CC   complex, abnormal activity of the TNF receptor complex, or abnormal
CC   activity of a caspase. Diseases that may be treated include cancer
CC   (particularly follicular lymphoma, carcinomas associated with mutations
CC   in p53 and hormone-dependent tumours), autoimmune disorders (e.g.
CC   systemic lupus erythematosis, immune-mediated glomerulonephritis), viral
CC   infections, Alzheimer's disease, Parkinson's disease, amyotrophic lateral
CC   sclerosis, retinitis pigmentosa, spinal muscular dystrophy, cerebellar
CC   degeneration, anaemia, myelodysplastic syndrome, myocardial infarction,
CC   and stroke. CARD-3 protein interacts with other cellular proteins, and so
CC   can be used for regulation of cellular proliferation and differentiation
CC   and cell survival. The CARD proteins may also be used to for screen drugs
CC   or compounds which modulate their activity. The CARD-4 gene can express a
CC   long transcript that encodes CARD-4L, a short transcript that encodes
CC   CARD-4S or two CARD-4 splice variants, CARD-4Y and CARD-4Z. This sequence
CC   represents the human CARD-3 protein described in the method of the
CC   invention.
XX
SQ   Sequence   540 AA;

  Query Match             99.6%;  Score 2817;  DB 20;  Length 540;
  Best Local Similarity   99.6%;  Pred. No. 6.9e-256;
  Matches  538;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 MNGEAICSALPTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSER 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mngeaicsalptipyhkladlrylsrgasgtvssarhadwrvqvavkhlhihtplldser 60

Qy      61 KDVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPL 120
           |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db      61 kdvlreaeilhkarfsyilpilgicnepeflgivteympngslnellhrkteypdvawpl 120

Qy     121 RFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 rfrilheialgvnylhnmtppllhhdlktqnilldnefhvkiadfglskwrmmslsqsrs 180

Qy     181 SKSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMY 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 sksapeggtiiymppenyepgqksrasikhdiysyavitwevlsrkqpfedvtnplqimy 240

Qy     241 SVSQGHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEI 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 svsqghrpvineeslpydiphrarmisliesgwaqnpderpsflkclielepvlrtfeei 300

Qy     301 TFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLNIPVNHGPQEESCGSSQLHENSGSPET 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 tfleaviqlkktklqsvssaihlcdkkkmelslnipvnhgpqeescgssqlhensgspet 360

Qy     361 SRSLPAPQDNDFLSRKAQDCYFMKLHHCPGNHSWDSTISGSQRAAFCDHKTTPCSSAIIN 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db     361 srslpapqdndflsrkaqdcyfmklhhcpgnhswdstisgsqraafcdhktipcssaiin 420

Qy     421 PLSTAGNSERLQPGIAQQWIQSKREDIVNQMTEACLNQSLDALLSRDLIMKEDYELVSTK 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 plstagnserlqpgiaqqwiqskredivnqmteaclnqsldallsrdlimkedyelvstk 480

Qy     481 PTRTSKVRQLLDTTDIQGEEFAKVIVQKLKDNKQMGLQPYPEILVVSRSPSLNLLQNKSM 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    481 ptrtskvrqlldttdiqgeefakvivqklkdnkqmglqpypeilvvsrspslnllqnksm 540

RESULT   5
Y59404
ID   Y59404 standard; Protein; 531 AA.
XX
AC   Y59404;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Human RICK protein sequence.
XX
KW   RICK; human; RIP-like interacting CLARP kinase; apoptosis regulator; ARC;
KW   caspase-8; caspase-10; CD95 signalling; apoptosis signalling pathway;
KW   CIDE-A; CIDE-B; DREP-1; diagnosis; cell growth; apoptosis dysregulation;
KW   gene therapy; cardiac disorder; cancer; neurodegenerative disorder; AIDS;
KW   aplastic anaemia; ischaemic injury; toxin-induced liver disease.
XX
OS   Homo sapiens.
XX
PN   WO9955134-A2.
XX
PD   04-NOV-1999.
XX
PF   27-APR-1999;   99WO-US09183.
XX
PR   27-APR-1998;   98US-0069023.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Nunez G,  Inohara N,  Koseki T;
XX
DR   WPI; 2000-072163/06.
DR   N-PSDB; Z48762.
XX
PT   Compositions for identifying apoptosis signalling pathway inhibitors
PT   useful for treating diseases -
XX
PS   Claim 1; Fig 7a; 93pp; English.
XX
CC   This sequence is the human RICK (RIP-like interacting CLARP kinase)
CC   protein of the invention. The RICK protein acts as a positive regulator
CC   of apoptosis, potentiating apoptosis induced by caspase-8 and caspase-10
CC   during CD95 signalling. The invention provides methods for identifying
CC   apoptosis signalling pathway inhibitors and activators, and methods and
CC   compositions for screening compounds which will modulate the interactions
CC   of the various compositions identified: ARC, RICK, and the CIDE family of
CC   activators (CIDE-A, CIDE-B and DREP-1). RICK is useful in screening
CC   assays for agents, useful in the diagnosis, prognosis or treatment of
CC   disease associated with excess cell growth and dysregulation of
CC   apoptosis. Complexes containing RICK and CLARP can be used in drug
CC   screening assays to identify inhibitor molecules blocking CD95-mediated
CC   apoptosis. Overexpression of ARC in an in vitro cell system can be used
CC   to identify inhibitors of the enzymatic activity of caspase-8.
CC   Identification of ARC-like inhibitory compounds may be useful for gene
CC   therapy treatment of disease with increased cell death in muscle tissue
CC   and cardiac disorders. Therapeutic compositions of CIDEs can be used to
CC   treat e.g. cancer, AIDS, neurodegenerative disorders, aplastic anaemia,
CC   ischaemic injury, and toxin-induced liver disease. AntiRICK antibodies
CC   can be used as reagents for the preparation or affinity chromatography
CC   media, and for diagnostically measuring RICK levels. A specific inhibitor
CC   of an essential step in the biochemistry of apoptosis is needed. RICK
CC   interaction with intracellular factors such as CLARP and FADD appears to
CC   be essential for apoptosis, inhibitors of RICK binding to intracellular
CC   apoptosis factors are potential drug candidates.
XX
SQ   Sequence   531 AA;

  Query Match             98.1%;  Score 2774;  DB 21;  Length 531;
  Best Local Similarity   99.6%;  Pred. No. 7.4e-252;
  Matches  529;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     10 LPTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSERKDVLREAEI 69
          :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 mptipyhkladlrylsrgasgtvssarhadwrvqvavkhlhihtplldserkdvlreaei 60

Qy     70 LHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILHEIA 129
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 lhkarfsyilpilgicnepeflgivteympngslnellhrkteypdvawplrfrilheia 120

Qy    130 LGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPEGGT 189
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 lgvnylhnmtppllhhdlktqnilldnefhvkiadfglskwrmmslsqsrssksapeggt 180

Qy    190 IIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVSQGHRPV 249
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 iiymppenyepgqksrasikhdiysyavitwevlsrkqpfedvtnplqimysvsqghrpv 240

Qy    250 INEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEITFLEAVIQL 309
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ineeslpydiphrarmisliesgwaqnpderpsflkclielepvlrtfeeitfleaviql 300

Qy    310 KKTKLQSVSSAIHLCDKKKMELSLNIPVNHGPQEESCGSSQLHENSGSPETSRSLPAPQD 369
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 kktklqsvssaihlcdkkkmelslnipvnhgpqeescgssqlhensgspetsrslpapqd 360

Qy    370 NDFLSRKAQDCYFMKLHHCPGNHSWDSTISGSQRAAFCDHKTTPCSSAIINPLSTAGNSE 429
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ndflsrkaqdcyfmklhhcpgnhswdstisgsqraafcdhkttpcssaiinplstagnse 420

Qy    430 RLQPGIAQQWIQSKREDIVNQMTEACLNQSLDALLSRDLIMKEDYELVSTKPTRTSKVRQ 489
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 rlqpgiaqqwiqskredivnqmteaclnqsldallsrdlimkedyelvstkptrtskvrq 480

Qy    490 LLDTTDIQGEEFAKVIVQKLKDNKQMGLQPYPEILVVSRSPSLNLLQNKSM 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 lldttdiqgeefakvivqklkdnkqmglqpypeilvvsrspslnllqnksm 531

RESULT   6
Y59405
ID   Y59405 standard; Protein; 478 AA.
XX
AC   Y59405;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Human RICK protein sequence residues 54-531.
XX
KW   RICK; human; RIP-like interacting CLARP kinase; apoptosis regulator; ARC;
KW   caspase-8; caspase-10; CD95 signalling; apoptosis signalling pathway;
KW   CIDE-A; CIDE-B; DREP-1; diagnosis; cell growth; apoptosis dysregulation;
KW   gene therapy; cardiac disorder; cancer; neurodegenerative disorder; AIDS;
KW   aplastic anaemia; ischaemic injury; toxin-induced liver disease.
XX
OS   Homo sapiens.
XX
PN   WO9955134-A2.
XX
PD   04-NOV-1999.
XX
PF   27-APR-1999;   99WO-US09183.
XX
PR   27-APR-1998;   98US-0069023.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Nunez G,  Inohara N,  Koseki T;
XX
DR   WPI; 2000-072163/06.
XX
PT   Compositions for identifying apoptosis signalling pathway inhibitors
PT   useful for treating diseases -
XX
```

```
PS   Claim 6; Page -; 93pp; English.
XX
CC   This sequence is a fragment of the human RICK (RIP-like interacting CLARP
CC   kinase) protein of the invention. RICK acts as a positive regulator
CC   of apoptosis, potentiating apoptosis induced by caspase-8 and caspase-10
CC   during CD95 signalling. The invention provides methods for identifying
CC   apoptosis signalling pathway inhibitors and activators, and methods and
CC   compositions for screening compounds which will modulate the interactions
CC   of the various compositions identified: ARC, RICK, and the CIDE family of
CC   activators (CIDE-A, CIDE-B and DREP-1). RICK is useful in screening
CC   assays for agents, useful in the diagnosis, prognosis or treatment of
CC   disease associated with excess cell growth and dysregulation of
CC   apoptosis. Complexes containing RICK and CLARP can be used in drug
CC   screening assays to identify inhibitor molecules blocking CD95-mediated
CC   apoptosis. Overexpression of ARC in an in vitro cell system can be used
CC   to identify inhibitors of the enzymatic activity of caspase-8.
CC   Identification of ARC-like inhibitory compounds may be useful for gene
CC   therapy treatment of disease with increased cell death in muscle tissue
CC   and cardiac disorders. Therapeutic compositions of CIDEs can be used to
CC   treat e.g. cancer, AIDS, neurodegenerative disorders, aplastic anaemia,
CC   ischaemic injury, and toxin-induced liver disease. AntiRICK antibodies
CC   can be used as reagents for the preparation or affinity chromatography
CC   media, and for diagnostically measuring RICK levels. A specific inhibitor
CC   of an essential step in the biochemistry of apoptosis is needed. RICK
CC   interaction with intracellular factors such as CLARP and FADD appears to
CC   be essential for apoptosis, inhibitors of RICK binding to intracellular
CC   apoptosis factors are potential drug candidates.
CC   Note: This sequence was created using information given in the
CC   specification.
XX
SQ   Sequence   478 AA;

  Query Match             88.3%;  Score 2497;  DB 21;  Length 478;
  Best Local Similarity   99.8%;  Pred. No. 6.7e-226;
  Matches  477;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         63 VLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRF 122
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db          1 vlreaeilhkarfsyilpilgicnepeflgivteympngslnellhrkteypdvawplrf 60

Qy        123 RILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSK 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 rilheialgvnylhnmtppllhhdlktqnilldnefhvkiadfglskwrmmslsqsrssk 120

Qy        183 SAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSV 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 sapeggtiiymppenyepgqksrasikhdiysyavitwevlsrkqpfedvtnplqimysv 180

Qy        243 SQGHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEITF 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 sqghrpvineeslpydiphrarmisliesgwaqnpderpsflkclielepvlrtfeeitf 240

Qy        303 LEAVIQLKKTKLQSVSSAIHLCDKKKMELSLNIPVNHGPQEESCGSSQLHENSGSPETSR 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 leaviqlkktklqsvssaihlcdkkkmelslnipvnhgpqeescgssqlhensgspetsr 300

Qy        363 SLPAPQDNDFLSRKAQDCYFMKLHHCPGNHSWDSTISGSQRAAFCDHKTTPCSSAIINPL 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 slpapqdndflsrkaqdcyfmklhhcpgnhswdstisgsqraafcdhkttpcssaiinpl 360

Qy        423 STAGNSERLQPGIAQQWIQSKREDIVNQMTEACLNQSLDALLSRDLIMKEDYELVSTKPT 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 stagnserlqpgiaqqwiqskredivnqmteaclnqsldallsrdlimkedyelvstkpt 420

Qy        483 RTSKVRQLLDTTDIQGEEFAKVIVQKLKDNKQMGLQPYPEILVVSRSPSLNLLQNKSM 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 rtskvrqlldttdiqgeefakvivqklkdnkqmglqpypeilvvsrspslnllqnksm 478

RESULT   7
Y59406
ID   Y59406 standard; Protein; 284 AA.
XX
AC   Y59406;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Human RICK protein sequence residues 248-531.
XX
KW   RICK; human; RIP-like interacting CLARP kinase; apoptosis regulator; ARC;
KW   caspase-8; caspase-10; CD95 signalling; apoptosis signalling pathway;
KW   CIDE-A; CIDE-B; DREP-1; diagnosis; cell growth; apoptosis dysregulation;
KW   gene therapy; cardiac disorder; cancer; neurodegenerative disorder; AIDS;
KW   aplastic anaemia; ischaemic injury; toxin-induced liver disease.
XX
OS   Homo sapiens.
XX
PN   WO9955134-A2.
XX
PD   04-NOV-1999.
XX
PF   27-APR-1999;   99WO-US09183.
XX
PR   27-APR-1998;   98US-0069023.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Nunez G,  Inohara N,  Koseki T;
XX
DR   WPI; 2000-072163/06.
XX
PT   Compositions for identifying apoptosis signalling pathway inhibitors
PT   useful for treating diseases -
XX
PS   Claim 6; Page -; 93pp; English.
XX
CC   This sequence is a fragment of the human RICK (RIP-like interacting CLARP
CC   kinase) protein of the invention. RICK acts as a positive regulator
CC   of apoptosis, potentiating apoptosis induced by caspase-8 and caspase-10
CC   during CD95 signalling. The invention provides methods for identifying
CC   apoptosis signalling pathway inhibitors and activators, and methods and
CC   compositions for screening compounds which will modulate the interactions
CC   of the various compositions identified: ARC, RICK, and the CIDE family of
CC   activators (CIDE-A, CIDE-B and DREP-1). RICK is useful in screening
CC   assays for agents, useful in the diagnosis, prognosis or treatment of
CC   disease associated with excess cell growth and dysregulation of
CC   apoptosis. Complexes containing RICK and CLARP can be used in drug
CC   screening assays to identify inhibitor molecules blocking CD95-mediated
CC   apoptosis. Overexpression of ARC in an in vitro cell system can be used
CC   to identify inhibitors of the enzymatic activity of caspase-8.
CC   Identification of ARC-like inhibitory compounds may be useful for gene
CC   therapy treatment of disease with increased cell death in muscle tissue
CC   and cardiac disorders. Therapeutic compositions of CIDEs can be used to
CC   treat e.g. cancer, AIDS, neurodegenerative disorders, aplastic anaemia,
CC   ischaemic injury, and toxin-induced liver disease. AntiRICK antibodies
CC   can be used as reagents for the preparation or affinity chromatography
CC   media, and for diagnostically measuring RICK levels. A specific inhibitor
CC   of an essential step in the biochemistry of apoptosis is needed. RICK
CC   interaction with intracellular factors such as CLARP and FADD appears to
CC   be essential for apoptosis, inhibitors of RICK binding to intracellular
CC   apoptosis factors are potential drug candidates.
CC   Note: This sequence was created using information given in the
CC   specification.
XX
SQ   Sequence   284 AA;

  Query Match             52.1%;  Score 1475;  DB 21;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-130;
  Matches  284;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        257 YDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEITFLEAVIQLKKTKLQS 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ydiphrarmisliesgwaqnpderpsflkclielepvlrtfeeitfleaviqlkktklqs 60
```

```
Qy    317 VSSAIHLCDKKKMELSLNIPVNHGPQEESCGSSQLHENSGSPETSRSLPAPQDNDFLSRK 376
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 vssaihlcdkkkmelslnipvnhgpqeescgssqlhensgspetsrslpapqdndflsrk 120

Qy    377 AQDCYFMKLHHCPGNHSWDSTISGSQRAAFCDHKTTPCSSAIINPLSTAGNSERLQPGIA 436
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 aqdcyfmklhhcpgnhswdstisgsqraafcdhkttpcssaiinplstagnserlqpgia 180

Qy    437 QQWIQSKREDIVNQMTEACLNQSLDALLSRDLIMKEDYELVSTKPTRTSKVRQLLDTTDI 496
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 qqwiqskredivnqmteaclnqsldallsrdlimkedyelvstkptrtskvrqlldttdi 240

Qy    497 QGEEFAKVIVQKLKDNKQMGLQPYPEILVVSRSPSLNLLQNKSM 540
          ||||||||||||||||||||||||||||||||||||||||||||
Db    241 qgeefakvivqklkdnkqmglqpypeilvvsrspslnllqnksm 284

RESULT   8
Y59407
ID   Y59407 standard; Protein; 167 AA.
XX
AC   Y59407;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Human RICK protein sequence residues 365-531.
XX
KW   RICK; human; RIP-like interacting CLARP kinase; apoptosis regulator; ARC;
KW   caspase-8; caspase-10; CD95 signalling; apoptosis signalling pathway;
KW   CIDE-A; CIDE-B; DREP-1; diagnosis; cell growth; apoptosis dysregulation;
KW   gene therapy; cardiac disorder; cancer; neurodegenerative disorder; AIDS;
KW   aplastic anaemia; ischaemic injury; toxin-induced liver disease.
XX
OS   Homo sapiens.
XX
PN   WO9955134-A2.
XX
PD   04-NOV-1999.
XX
PF   27-APR-1999;   99WO-US09183.
XX
PR   27-APR-1998;   98US-0069023.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Nunez G,  Inohara N,  Koseki T;
XX
DR   WPI; 2000-072163/06.
XX
PT   Compositions for identifying apoptosis signalling pathway inhibitors
PT   useful for treating diseases -
XX
PS   Claim 6; Page -; 93pp; English.
XX
CC   This sequence is a fragment of the human RICK (RIP-like interacting CLARP
CC   kinase) protein of the invention. RICK acts as a positive regulator
CC   of apoptosis, potentiating apoptosis induced by caspase-8 and caspase-10
CC   during CD95 signalling. The invention provides methods for identifying
CC   apoptosis signalling pathway inhibitors and activators, and methods and
CC   compositions for screening compounds which will modulate the interactions
CC   of the various compositions identified: ARC, RICK, and the CIDE family of
CC   activators (CIDE-A, CIDE-B and DREP-1). RICK is useful in screening
CC   assays for agents, useful in the diagnosis, prognosis or treatment of
CC   disease associated with excess cell growth and dysregulation of
CC   apoptosis. Complexes containing RICK and CLARP can be used in drug
CC   screening assays to identify inhibitor molecules blocking CD95-mediated
CC   apoptosis. Overexpression of ARC in an in vitro cell system can be used
CC   to identify inhibitors of the enzymatic activity of caspase-8.
CC   Identification of ARC-like inhibitory compounds may be useful for gene
CC   therapy treatment of disease with increased cell death in muscle tissue
CC   and cardiac disorders. Therapeutic compositions of CIDEs can be used to
CC   treat e.g. cancer, AIDS, neurodegenerative disorders, aplastic anaemia,
CC   ischaemic injury, and toxin-induced liver disease. AntiRICK antibodies
CC   can be used as reagents for the preparation or affinity chromatography
CC   media, and for diagnostically measuring RICK levels. A specific inhibitor
CC   of an essential step in the biochemistry of apoptosis is needed. RICK
CC   interaction with intracellular factors such as CLARP and FADD appears to
CC   be essential for apoptosis, inhibitors of RICK binding to intracellular
CC   apoptosis factors are potential drug candidates.
CC   Note: This sequence was created using information given in the
CC   specification.
XX
SQ   Sequence   167 AA;

  Query Match             30.6%;  Score 867;  DB 21;  Length 167;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-73;
  Matches  167;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    374 SRKAQDCYFMKLHHCPGNHSWDSTISGSQRAAFCDHKTTPCSSAIINPLSTAGNSERLQP 433
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 srkaqdcyfmklhhcpgnhswdstisgsqraafcdhkttpcssaiinplstagnserlqp 60

Qy    434 GIAQQWIQSKREDIVNQMTEACLNQSLDALLSRDLIMKEDYELVSTKPTRTSKVRQLLDT 493
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 giaqqwiqskredivnqmteaclnqsldallsrdlimkedyelvstkptrtskvrqlldt 120

Qy    494 TDIQGEEFAKVIVQKLKDNKQMGLQPYPEILVVSRSPSLNLLQNKSM 540
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    121 tdiqgeefakvivqklkdnkqmglqpypeilvvsrspslnllqnksm 167

RESULT   9
B58938
ID   B58938 standard; Protein; 153 AA.
XX
AC   B58938;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Breast and ovarian cancer associated antigen protein sequence SEQ ID 646.
XX
KW   Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW   nootropic; neurpprotective; antiviral; antiallergic; hepatotropic;
KW   antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;
KW   antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
KW   Addison's disease; allergy; autoimmune haemolytic anaemia;
KW   autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW   multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW   cardiovascular disorder; wound healing; neurological disease.
XX
OS   Homo sapiens.
XX
PN   WO200055173-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05881.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-611515/58.
DR   N-PSDB; F21841.
XX
PT   New human breast and ovarian cancer associated gene sequences and the
PT   polypeptides encoded by these genes, useful in the prevention,
PT   treatment and diagnosis of cancer, immune disorders, cardiovascular
PT   disorders and neurological diseases -
XX
PS   Claim 11; Page 1086; 1299pp; English.
XX
CC   Sequences F21614 - F22031 represent DNA sequences encoding human proteins
```

```
CC   B58711 - B59128. The DNA and protein sequences are associated with
CC   breast and ovarian cancer. Included in the invention are sequences
CC   F22032 - F22040 and B59129 which are used in the isolation and
CC   characterisation of the DNA and protein sequences of the invention. The
CC   breast and ovarian cancer associated DNA, protein, agonist or antagonist
CC   sequences exhibit cytostatic; immunosuppressive; nootropic;
CC   neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
CC   antiinflammatory; antiulcer; vulnerary; anticonvulsant; antibacterial;
CC   antifungal; antiparasitic and cardiant activity. The polynucleotide and
CC   protein sequences are used in the diagnosis of cancer, particularly
CC   breast and ovarian cancer. The nucleic acid sequences, proteins, agonists
CC   and agonists may also be used in the diagnosis, prevention and treatment
CC   of immune disorders e.g. Addison's disease, allergies, autoimmune
CC   haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC   disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC   cardiovascular disorders such as myocardial ischaemias; wound healing;
CC   neurological diseases such as cerebral anoxia and epilepsy; and
CC   infectious diseases.
XX
SQ   Sequence   153 AA;

  Query Match             28.2%;  Score 797;  DB 21;  Length 153;
  Best Local Similarity   99.3%;  Pred. No. 5.9e-67;
  Matches  149;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     97 YMPNGSLNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDN 156
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ympngslnellhrkteypdvawplrfrilheialgvnylhnmtppllhhdlktqnilldn 60

Qy    157 EFHVKIADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYA 216
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 efhvkiadfglskwrmmslsqsrssksapeggtiiymppenyepgqksrasikhdiysya 120

Qy    217 VITWEVLSRKQPFEDVTNPLQIMYSVSQGH 246
          |||||| |||||||||||||||||||||||
Db    121 vitwevxsrkqpfedvtnplqimysvsqgh 150

RESULT  10
Y69163
ID   Y69163 standard; Protein; 786 AA.
XX
AC   Y69163;
XX
DT   30-MAY-2000  (first entry)
XX
DE   Amino acid sequence of a death associated kinase with ankyrin repeats.
XX
KW   Death associated kinase protein containing ankyrin repeats; DAKAR;
KW   kinase; quality assurance agent; shelf life; marker.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "Met encoded by GTG"
FT   Domain          17..297
FT                   /note= "N-terminal kinase domain"
FT   Region          471..768
FT                   /note= "this region contains a series of nine
FT                          tandem ankyrin repeats"
XX
PN   WO200008177-A2.
XX
PD   17-FEB-2000.
XX
PF   04-AUG-1999;   99WO-US17576.
XX
PR   04-AUG-1998;   98US-0095269.
PR   11-SEP-1998;   98US-0099973.
PR   09-FEB-1999;   99US-0119353.
XX
PA   (IMMV ) IMMUNEX CORP.
XX
PI   Bird TA,  Virca GD;
XX
DR   WPI; 2000-195582/17.
DR   N-PSDB; Z61161.
XX
PT   Novel death associated kinase containing ankyrin repeats (DAKAR) used
PT   as molecular weight marker and as controls for peptide fragmentation  -
XX
PS   Claim 13; Page 10; 71pp; English.
XX
CC   The present sequence represents a murine death associated kinase protein,
CC   containing ankyrin repeats, designated DAKAR. The DAKAR polynucleotides
CC   can be used to express the polypeptides, and as probes to identify
CC   nucleic acids encoding proteins having kinase activity. DAKAR
CC   polypeptides and fragmented polypeptides are used for purifying
CC   proteins, e.g. to measure protein activity; as quality assurance agents
CC   to monitor shelf life and stability of binding partner proteins; as
CC   research agents, e.g. in assays to determine protein kinase activity,
CC   to identify novel molecules involved in signal transduction pathways,
CC   and to identify therapeutic compounds which may interfere with
CC   apoptosis; as molecular weight and isoelectric focusing markers; as
CC   controls for peptide fragmentation; identification of unknown proteins,
CC   e.g. by comparison with proteins in databases; and for preparation of
CC   antibodies. The antibodies can be used in assays to detect the presence
CC   of the protein, and to purify the protein by immunoaffinity
CC   chromatography. The antibodies can also be used to block binding of
CC   the DAKAR polypeptides to their binding partners. Compounds that inhibit
CC   or enhance the kinase activity of DAKAR can be used to treat diseases
CC   characterized by overproduction or upregulated production or
CC   underproduction or downregulated production of DAKAR.
XX
SQ   Sequence   786 AA;

  Query Match             20.4%;  Score 577;  DB 21;  Length 786;
  Best Local Similarity   33.3%;  Pred. No. 3.7e-45;
  Matches  172;  Conservative   64;  Mismatches  160;  Indels  120;  Gaps   21;

Qy      1 MNGEA----ICSALPTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVK---HLHIHT 53
          | ||         | |    : |    :  |  | |   ||  |:  :|:|    ||:
Db      1 megegrgrwalgllrtfdagefagwekvgsggfgqvykvrhvhwktwlaikcspslhvd- 59

Qy     54 PLLDSERKDVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEY 113
             | || ::| ||: :  |:| || |: ||| ||  :|:| |||  ||| :||  :
Db     60 ---drermelleeakkmemakfryilpvygicqep--vglvmeymetgslekllase--- 111

Qy    114 PDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMM 173
            : | |||||:|| |:|:|:|| |:||||| |||  |||||  :||||:||||:|   |
Db    112 -plpwdlrfrivhetavgmnflhcmspplhldlkpanilldahyhvkisdfglakcngm 170

Qy    174 SLSQSRSSKSAPEG--GTIIYMPPENYEPGQKSRA-SIKHDIYSYAVITWEVLSRKQPFE 230
          | |   |     :|  ||| |:|||     :|||    |||:||:|:: | ||::|:||
Db    171 shshdlsm----dglfgtiaylpperir--eksrlfdtkhdvysfaiviwgvltqkkpfa 224

Qy    231 DVTNPLQIMYSVSQGHRPVINEESLPYDIPHR---ARMISLIESGWAQNPDERPSFLKCL 287
          |  | | ||  | :||||    |  |   |     | :| |::  |  :|  ||
Db    225 deknilhimmkvvkghrp----elppicrprpracasliglmqrcwhadpqvrp------ 274

Qy    288 IELEPVLRTFEEITFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLNIPVNHGPQEESCG 347
                  ||:|||                |    ||:|   |:     : | | |:|
Db    275 --------tfqeit----------------setedlcekpdeevk---dlahepgekssl 307

Qy    348 SSQLHENSGSPETSRSLPAPQDNDF----------------------LSRKAQDCYFMKL 385
           |:      |    |:   | |||                       ||| : :|   ||
Db    308 esksearpessrlkrasappfdndcslsellsqldsgisqtlegpeelsrsssec---kl 364

Qy    386 ------HHCPGNHSWDSTIS--GSQRAAF-----------CDHKTTPCSSAIINPLSTAG 426
                    |  | ||  |  ||   :|            | :      |||     :|
Db    365 pssssgkrlsgvssvdsafssrgslslsfereastgdlgptdiqkkklvdaii-----sg 419

Qy    427 NSERLQPGIAQQWIQSKREDIVNQMTEACLNQSLDA 462
```

```
          :: ||      : :| :  |:|   : : |: :::|
Db    420 dtsrl-----mkilqpqdvdlvldssasllhlavea 450


RESULT  11
Y76079
ID   Y76079 standard; Protein; 787 AA.
XX
AC   Y76079;
XX
DT   27-MAR-2000  (first entry)
XX
DE   Murine protein kinase/ankyrin homologue, SEQ ID NO:334.
XX
KW   Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
KW   embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
KW   secreted; transmembrane; inflammation; cancer; neurological disease;
KW   angiogenesis; tumour vascularisation; growth disorder;
KW   developmental disorder; skin wound; hair follicle disorder;
KW   anti-inflammatory; cytostatic; neuroprotective; vulnery.
XX
OS   Mus sp.
XX
PN   WO9955865-A1.
XX
PD   04-NOV-1999.
XX
PF   29-APR-1999;   99WO-NZ00051.
XX
PR   29-APR-1998;   98US-0069726.
PR   09-NOV-1998;   98US-0188930.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Strachan L,  Sleeman M,  Watson JD,  Onrust R,  Kumble A,  Murison JG;
XX
DR   WPI; 2000-072177/06.
DR   N-PSDB; Z61784.
XX
PT   Novel polynucleotides useful for the treatment of various conditions
PT   including wounds and cancer -
XX
PS   Claim 4; Page 195-196; 235pp; English.
XX
CC   The invention relates to novel nucleic acid sequences derived from rat
CC   dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
CC   and mouse embryonic skin, keratinocyte stem cells and transit amplifying
CC   cells. Polypeptides of the invention may be used to treat inflammation,
CC   cancer and neurological diseases. The proteins may be used to stimulate
CC   the growth and motility of keratinocytes, to inhibit the growth of
CC   cancer cells, to modulate angiogenesis and tumour vascularisation, to
CC   modulate skin inflammation, to modulate epithelial cell growth and to
CC   inhibit binding of HIV-1 to leukocytes. The invention may also be used to
CC   treat growth and developmental defects, skin wounds and hair follicle
CC   disorders. Sequences Y75942-Y76123 represent polypeptides encoded
CC   by cDNA sequences derived from several mouse, rat or human skin cell
CC   types. Sequences Y75942-Y75947, Y76020-Y76021, Y76094-Y76104 and
CC   Y76119 are proteins with an N-terminal signal sequence, indicating that
CC   they are secreted. Sequences Y75986-Y75989, Y76061-Y76071, Y76106-Y76109
CC   and Y76121-Y76122 are proteins with one or more putative transmembrane
CC   domains.
XX
SQ   Sequence   787 AA;


  Query Match            20.4%;  Score 576;  DB 21;  Length 787;
  Best Local Similarity   33.6%;  Pred. No. 4.6e-45;
  Matches  169;  Conservative   64;  Mismatches  154;  Indels  116;  Gaps   20;

Qy     10 LPTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVK---HLHIHTPLLDSERKDVLRE 66
          | |    : |    :  |  | |   ||  |:  :|:|    ||:     | || ::| |
Db     15 lrtfdagefagwekvgsggfgqvykvrhvhwktwlaikcspslhvd---drermellee 70

Qy     67 AEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILH 126
          |: :  |:| || |: ||| ||  :|:| |||  ||| :||  :     : | |||||:|
Db     71 akkmemakfryilpvygicqep--vglvmeymetgslekllase----plpwdlrfrivh 124

Qy    127 EIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPE 186
          | |:|:|:|| |:||||| |||  |||||  :||||:||||:|   || |   |     :
Db    125 etavgmnflhcmspllhldlkpanilldahyhvkisdfglakcngmshshdlsm----d 180

Qy    187 G--GTIIYMPPENYEPGQKSRA-SIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVS 243
          |  ||| |:|||     :|||    |||:||:|:: | ||::|:|| |  | | ||  |
Db    181 glfgtiaylpperir--eksrlfdtkhdvysfaiviwgvltqkkpfadeknilhimmkvv 238

Qy    244 QGHRPVINEESLPYDIPHR---ARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEI 300
          :||||    |  |   |     | :| |::  |  :|  ||              ||:||
Db    239 kghrp----elppicrprpracasliglmqrcwhadpqvrp-------------tfqei 280

Qy    301 TFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLNIPVNHGPQEESCGSSQLHENSGSPET 360
          |                |    ||:|   |:     : | | |:|   |:      |
Db    281 t---------------setedlcekpdeevk---dlahepgeksslesksearpessrl 321

Qy    361 SRSLPAPQDNDF----------------------LSRKAQDCYFMKL------HHCPGNH 392
           |:   | |||                       ||| : :|   ||          |
Db    322 krasappfdndcslsellsqldsgisqtlegpeelsrsssec---klpssssgkrlsgvs 378

Qy    393 SWDSTIS--GSQRAAF-----------CDHKTTPCSSAIINPLSTAGNSERLQPGIAQQW 439
          | ||  |  ||   :|            | :      |||     :|:: ||      :
Db    379 svdsafssrgslslsfereastgdlgptdiqkkklvdaii----sgdtsrl-----mki 428

Qy    440 IQSKREDIVNQMTEACLNQSLDA 462
          :| :  |:|   : : |: :::|
Db    429 lqpqdvdlvldssasllhlavea 451


RESULT  12
B56018
ID   B56018 standard; Protein; 787 AA.
XX
AC   B56018;
XX
DT   08-MAR-2001  (first entry)
XX
DE   Skin cell protein, SEQ ID NO: 334.
XX
KW   Mouse; skin cell; cytostatic; antiinflammatory; anti-HIV;
KW   nootropic; neuroprotective; vulnerary; immunomodulatory; vaccine;
KW   keratinocyte growth stimulation; cancer; angiogenesis inhibition;
KW   inflammation; neurological disease.
XX
OS   Mus sp.
XX
PN   WO200069884-A2.
XX
PD   23-NOV-2000.
XX
PF   15-MAY-2000; 2000WO-NZ00075.
XX
PR   14-MAY-1999;   99US-0312283.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Watson JD,  Strachan L,  Onrust R,  Sleeman M,  Kumble KD,  Murison JG;
XX
DR   WPI; 2001-007495/01.
DR   N-PSDB; C99717.
XX
PT   New isolated polynucleotide used in the identification of genetic
PT   disorders and encoding polypeptides used for treating inflammatory
PT   disease, cancer and neurological diseases -
XX
PS   Claim 4; Page 262-263; 352pp; English.
XX
CC   The present sequence is a polypeptide which is expressed in
CC   mammalian skin cells. The polypeptide is useful for stimulating
```

```
CC   keratinocyte growth and motility, inhibiting the growth of cancer cells,
CC   modulating angiogenesis, inhibiting angiogenesis and vascularisation of
CC   tumours, modulating skin inflammation, stimulating the growth of
CC   epithelial cells, inhibiting the binding of human immunodeficiency virus
CC   (HIV)-1 to leukocytes, and treating inflammmatory disease, cancer and
CC   neurological diseases. The polynucleotide can be used as a marker, in
CC   the identification of genetic disorders, and for the design of
CC   oligonucleotides for examining expression patterns.
XX
SQ   Sequence   787 AA;

  Query Match             20.4%;  Score 576;  DB 22;  Length 787;
  Best Local Similarity   33.6%;  Pred. No. 4.6e-45;
  Matches  169;  Conservative   64;  Mismatches  154;  Indels  116;  Gaps   20;

Qy      10 LPTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVK---HLHIHTPLLDSERKDVLRE 66
           | |    : |    :  |  | |   ||  |:  :|:|    ||:     | || ::| |
Db      15 lrtfdagefagwekvgsggfgqvykvrhvhwktwlaikcspslhvd---drermellee 70

Qy      67 AEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILH 126
           |: :  |:| || |: ||| ||  :|:| |||  ||| :||  :     : | |||||:|
Db      71 akkmemakfryilpvygicqep--vglvmeymetgsleklase----plpwdlrfrivh 124

Qy     127 EIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPE 186
           | |:|:|:|| |:||||| |||  |||||  :||||:||||:|   || |   |     :
Db     125 etavgmnflhcmspplhldlkpanilldahyhvkisdfglakcngmshshdlsm----d 180

Qy     187 G--GTIIYMPPENYEPGQKSRA-SIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVS 243
           |  ||| |:|||     :|||    |||:||:|:: | ||::|:|| |  | | ||  |
Db     181 glfgtiaylpperir--eksrlfdtkhdvysfaiviwgvltqkkpfadeknilhimmkvv 238

Qy     244 QGHRPVINEESLPYDIPHR---ARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEI 300
           :||||    |  |   |     | :| |::  |  :|  ||              ||:||
Db     239 kghrp----elppicrprpracasliglmqrcwhadpqvrp--------------tfqei 280

Qy     301 TFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLNIPVNHGPQEESCGSSQLHENSGSPET 360
           |                |    ||:|   |:     : | | |:|   |:      |
Db     281 t----------------setedlcekpdeevk---dlahepgekssleskse arpessrl 321

Qy     361 SRSLPAPQDNDF----------------------LSRKAQDCYFMKL------HHCPGNH 392
            |:   | |||                       ||| : :|   ||          |
Db     322 krasappfdndcslsellsqldsgisqtlegpeelsrsssec---klpssssgkrlsgvs 378

Qy     393 SWDSTIS--GSQRAAF-----------CDHKTTPCSSAIINPLSTAGNSERLQPGIAQQW 439
           | ||  |  ||   :|            | :      |||     :|:: ||      :
Db     379 svdsafssrgslslsfereastgdlgptdiqkkklvdaii-----sgdtsrl-----mki 428

Qy     440 IQSKREDIVNQMTEACLNQSLDA 462
           :| :  |:|   : : |: :::|
Db     429 lqpqdvdlvldssasllhlavea 451


RESULT  13
Y76123
ID   Y76123 standard; Protein; 590 AA.
XX
AC   Y76123;
XX
DT   27-MAR-2000  (first entry)
XX
DE   Murine RIP protein kinase homologue, SEQ ID NO:409.
XX
KW   Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
KW   embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
KW   secreted; transmembrane; inflammation; cancer; neurological disease;
KW   angiogenesis; tumour vascularisation; growth disorder;
KW   developmental disorder; skin wound; hair follicle disorder;
KW   anti-inflammatory; cytostatic; neuroprotective; vulnery.
XX
OS   Mus sp.
XX
PN   WO9955865-A1.
XX
PD   04-NOV-1999.
XX
PF   29-APR-1999;   99WO-NZ00051.
XX
PR   29-APR-1998;   98US-0069726.
PR   09-NOV-1998;   98US-0188930.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Strachan L,  Sleeman M,  Watson JD,  Onrust R,  Kumble A,  Murison JG;
XX
DR   WPI; 2000-072177/06.
DR   N-PSDB; Z61830.
XX
PT   Novel polynucleotides useful for the treatment of various conditions
PT   including wounds and cancer -
XX
PS   Claim 4; Page 231-232; 235pp; English.
XX
CC   The invention relates to novel nucleic acid sequences derived from rat
CC   dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
CC   and mouse embryonic skin, keratinocyte stem cells and transit amplifying
CC   cells. Polypeptides of the invention may be used to treat inflammation,
CC   cancer and neurological diseases. The proteins may be used to stimulate
CC   the growth and motility of keratinocytes, to inhibit the growth of
CC   cancer cells, to modulate angiogenesis and tumour vascularisation, to
CC   modulate skin inflammation, to modulate epithelial cell growth and to
CC   inhibit binding of HIV-1 to leukocytes. The invention may also be used to
CC   treat growth and developmental defects, skin wounds and hair follicle
CC   disorders. Sequences Y75942-Y76123 represent polypeptides encoded
CC   by cDNA sequences derived from several mouse, rat or human skin cell
CC   types. Sequences Y75942-Y75947, Y76020-Y76021, Y76094-Y76104 and
CC   Y76119 are proteins with an N-terminal signal sequence, indicating that
CC   they are secreted. Sequences Y75986-Y75989, Y76061-Y76071, Y76106-Y76109
CC   and Y76121-Y76122 are proteins with one or more putative transmembrane
CC   domains.
XX
SQ   Sequence   590 AA;

  Query Match             20.3%;  Score 574;  DB 21;  Length 590;
  Best Local Similarity   33.4%;  Pred. No. 4.5e-45;
  Matches  168;  Conservative   65;  Mismatches  154;  Indels  116;  Gaps   20;

Qy      10 LPTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVK---HLHIHTPLLDSERKDVLRE 66
           | |    : |    :  |  | |   ||  |:  :|:|    ||:     | || ::| |
Db      15 lrtfdagefagwekvgsggfgqvykvrhvhwktwlaikcspslhvd---drermellee 70

Qy      67 AEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILH 126
           |: :  |:| || |: ||| ||  :|:| |||  ||| :||  :     : | |||||:|
Db      71 akkmemakfryilpvygicqep--vglvmeymetgsleklase----plpwdlrfrivh 124

Qy     127 EIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPE 186
           | |:|:|:|| |:||||| |||  |||||  :||||:||||:|   || |   |     :
Db     125 etavgmnflhcmspplhldlkpanilldahyhvkisdfglakcngmshshdlsm----d 180

Qy     187 G--GTIIYMPPENYEPGQKSRA-SIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVS 243
           |  ||| |:|||     :|||    |||:||:|:: | ||::|:|| |  | | ||  |
Db     181 glfgtiaylpperir--eksrlfdtkhdvysfaiviwgvltqkkpfadeknilhimmkvv 238

Qy     244 QGHRPVINEESLPYDIPHR---ARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEI 300
           :||||    |  |   |     | :| :::  |  :|  ||              ||:||
Db     239 kghrp----elppicrprpracasligimqrcwhadpqvrp--------------tfqei 280

Qy     301 TFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLNIPVNHGPQEESCGSSQLHENSGSPET 360
           |                |    ||:|   |:     : | | |:|   |:      |
Db     281 t----------------setedlcekpdeevk---dlahepgekssleskse arpessrl 321

Qy     361 SRSLPAPQDNDF----------------------LSRKAQDCYFMKL------HHCPGNH 392
            |:   | |||                       ||| : :|   ||          |
Db     322 krasappfdndcslsellsqldsgisqtlegpeelsrsssec---klpssssgkrlsgvs 378
```

```
Qy    393 SWDSTIS--GSQRAAF-----------CDHKTTPCSSAIINPLSTAGNSERLQPGIAQQW 439
          | ||  |  ||   :|            | :      |||     :|:: ||      :
Db    379 svdsafssrgslslsfereastgdlgptdiqkkklvdaii-----sgdtsrl-----mki 428

Qy    440 IQSKREDIVNQMTEACLNQSLDA 462
          :| :  |:|   : : |: :::|
Db    429 lqpqdvdlvldssasllhlavea 451


RESULT  14
B56062
ID   B56062 standard; Protein; 590 AA.
XX
AC   B56062;
XX
DT   08-MAR-2001  (first entry)
XX
DE   Skin cell protein, SEQ ID NO: 409.
XX
KW   Mouse; skin cell; cytostatic; antiinflammatory; anti-HIV;
KW   nootropic; neuroprotective; vulnerary; immunomodulatory; vaccine;
KW   keratinocyte growth stimulation; cancer; angiogenesis inhibition;
KW   inflammation; neurological disease.
XX
OS   Mus sp.
XX
PN   WO200069884-A2.
XX
PD   23-NOV-2000.
XX
PF   15-MAY-2000; 2000WO-NZ00075.
XX
PR   14-MAY-1999;   99US-0312283.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Watson JD,  Strachan L,  Onrust R,  Sleeman M,  Kumble KD,  Murison JG;
XX
DR   WPI; 2001-007495/01.
DR   N-PSDB; C99763.
XX
PT   New isolated polynucleotide used in the identification of genetic
PT   disorders and encoding polypeptides used for treating inflammatory
PT   disease, cancer and neurological diseases -
XX
PS   Claim 4; Page 309-310; 352pp; English.
XX
CC   The present sequence is a polypeptide which is expressed in
CC   mammalian skin cells. The polypeptide is useful for stimulating
CC   keratinocyte growth and motility, inhibiting the growth of cancer cells,
CC   modulating angiogenesis, inhibiting angiogenesis and vascularisation of
CC   tumours, modulating skin inflammation, stimulating the growth of
CC   epithelial cells, inhibiting the binding of human immunodeficiency virus
CC   (HIV)-1 to leukocytes, and treating inflammmatory disease, cancer and
CC   neurological diseases. The polynucleotide can be used as a marker, in
CC   the identification of genetic disorders, and for the design of
CC   oligonucleotides for examining expression patterns.
XX
SQ   Sequence   590 AA;

  Query Match             20.3%;  Score 574;  DB 22;  Length 590;
  Best Local Similarity   33.4%;  Pred. No. 4.5e-45;
  Matches  168;  Conservative   65;  Mismatches  154;  Indels  116;  Gaps   20;

Qy     10 LPTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVK---HLHIHTPLLDSERKDVLRE 66
          | |    : |    :  |  | |   ||  | : :|:|    ||:     | || ::| |
Db     15 lrtfdagefagwekvgsggfgqvykvrhvhwktwlaikcspslhvd----drermellee 70

Qy     67 AEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILH 126
          |: :  |:| || |: ||| ||  :|:| |||  ||| :||  :     : | |||||:|
Db     71 akkmemakfryilpvygicqep--vglvmeymetgslekllase----plpwdlrfrivh 124
```

```
Qy    127 EIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPE 186
          |.|:|:|:|| |:||||| |||  |||||  :||||:||||:|   || |   |     :
Db    125 etavgmnflhcmspllhldlkpanilldahyhvkisdfglakcngmshshdlsm----d 180

Qy    187 G--GTIIYMPPENYEPGQKSRA-SIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVS 243
          |  ||| |:|||     :|||    |||:||:|:: | ||::|:|| |  | | ||  |
Db    181 glfgtiaylpperir--eksrlfdtkhdvysfaiviwgvltqkkpfadeknilhimmkvv 238

Qy    244 QGHRPVINEESLPYDIPHR---ARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEI 300
          :||||    |  |   |     | :| :::  |  :|  ||              ||:||
Db    239 kghrp----elppicrprpracasligimqrcwhadpqvrp--------------tfqei 280

Qy    301 TFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLNIPVNHGPQEESCGSSQLHENSGSPET 360
          |                |    ||:|   |:     : | | |:|   |:      |
Db    281 t----------------setedlcekpdeevk---dlahepgekssleksearpessrl 321

Qy    361 SRSLPAPQDNDF----------------------LSRKAQDCYFMKL------HHCPGNH 392
           |:   | |||                       ||| : :|   ||          |
Db    322 krasappfdndcslsellsqldsgisqtlegpeelsrsssec---klpssssgkrlsgvs 378

Qy    393 SWDSTIS--GSQRAAF-----------CDHKTTPCSSAIINPLSTAGNSERLQPGIAQQW 439
          | ||  |  ||   :|            | :      |||     :|:: ||      :
Db    379 svdsafssrgslslsfereastgdlgptdiqkkklvdaii-----sgdtsrl-----mki 428

Qy    440 IQSKREDIVNQMTEACLNQSLDA 462
          :| :  |:|   : : |: :::|
Db    429 lqpqdvdlvldssasllhlavea 451


RESULT  15
Y79154
ID   Y79154 standard; Protein; 763 AA.
XX
AC   Y79154;
XX
DT   05-JUN-2000  (first entry)
XX
DE   Mouse protein kinase Fel.
XX
KW   Fel; protein kinase; mouse; signal transduction.
XX
OS   Mus musculus.
XX
PN   WO200008178-A2.
XX
PD   17-FEB-2000.
XX
PF   04-AUG-1999;   99WO-US17577.
XX
PR   04-AUG-1998;   98US-0095269.
PR   11-SEP-1998;   98US-0099973.
XX
PA   (IMMV ) IMMUNEX CORP.
XX
PI   Virca GD,  Bird TA,  Anderson DM,  Marken JS;
XX
DR   WPI; 2000-205722/18.
DR   N-PSDB; Z58584.
XX
PT   Novel murine polynucleotides encoding kinase polypeptides, used as
PT   probes to identify nucleic acids encoding proteins having kinase
PT   activity -
XX
PS   Claim 2; Fig 7; 93pp; English.
XX
CC   The present sequence is that of murine Fel, a novel protein kinase.
CC   The sequence represents translation (reading frame 3) of Genesis
CC   clone 971025TRAM004820HT (see Z58584).  The invention relates to
CC   purified murine polypeptides having kinase function (see Y79152-56)
CC   and isolated nucleic acids encoding them (see Z58582-86).  Claimed
CC   vectors comprising the novel nucleic acids, and claimed host cells
CC   (bacterial, yeast, plant, insect or animal) transfected or
CC   transduced with the vectors, are used to produce the polypeptides.
```

```
CC   The kinase polypeptides and their fragments are used as mol.wt. and
CC   isoelectric focusing markers, and as controls for peptide
CC   fragmentation.  They also have a number of therapeutic uses, as
CC   kinases play a central role in cellular signal transduction.  As
CC   such, alterations in kinase expression and/or activation can have a
CC   profound effect on a plethora of cellular processes, e.g.
CC   proliferation and programmed cell death (apoptosis).  The kinase
CC   polypeptides could also be used to identify binding partner
CC   proteins, or as reagents to identify proteins that regulate or
CC   interact with them.  They may also be used for preparation of
CC   antibodies useful for detection, purification, and for blocking
CC   binding of kinase polypeptides to their binding partners.
XX
SQ   Sequence   763 AA;

  Query Match             20.3%;  Score 574;  DB 21;  Length 763;
  Best Local Similarity   34.2%;  Pred. No. 6.8e-45;
  Matches  166;  Conservative   62;  Mismatches  142;  Indels  116;  Gaps   20;

Qy       27 GASGTVSSARHADWRVQVAVK---HLHIHTPLLDSERKDVLREAEILHKARFSYIFPILG 83
            |  | |   ||  |:  :|:|    ||:     | || ::| ||: :  |:| || |: |
Db        8 ggfgqvykvrhvhwktwlaikcspslhvd----drermelleeakkmemakfryilpvyg 63

Qy       84 ICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLL 143
            || ||  :|:| |||  ||| :||  :     : | |||||:|| |:|:|:|| |:||||
Db       64 icqep--vglvmeymetgslekllase----plpwdlrfrivhetavgmnflhcmsppll 117

Qy      144 HHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPEG--GTIIYMPPENYEPG 201
            | |||  |||||  :||||:||||:|   || |  .|     :|  ||| |:|||
Db      118 hldlkpanilldahyhvkisdfglakcngmshshdlsm----dglfgtiaylpperir-- 171

Qy      202 QKSRA-SIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIP 260
            :|||    |||:||:|:: | ||::|:|| |  | | ||  | :||||    |  |   |
Db      172 eksrlfdtkhdvysfaiviwgvltqkkpfadeknilhimmkvvkghrp----elppicrp 227

Qy      261 HR---ARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEITFLEAVIQLKKTKLQSV 317
                 | :| |::  |  :|  ||              ||:|||
Db      228 rpracasliglmqrcwhadpqvrp--------------tfqeit---------------- 257

Qy      318 SSAIHLCDKKKMELSLNIPVNHGPQEESCGSSQLHENSGSPETSRSLPAPQDNDF----- 372
            |    ||:|   |:     : | | |:|   |:      |    |:   | |||
Db      258 setedlcekpdeevk---dlahepgeksslesksearpessrlkrasappfdndcslsel 314

Qy      373 -----------------LSRKAQDCYFMKL------HHCPGNHSWDSTIS--GSQRAAF- 406
                             ||| : :|   ||          |  | ||  |  ||   :|
Db      315 lsqldsgisqtlegpeelsrsssec---klpssssgkrlsgvssvdsafssrgslslsfe 371

Qy      407 ----------CDHKTTPCSSAIINPLSTAGNSERLQPGIAQQWIQSKREDIVNQMTEACL 456
                       | :      |||     :|:: ||      : :| :  |:|   : : |
Db      372 reastgdlgptdiqkkklvdaii-----sgdtsrl-----mkilqpqdvdlvldssasll 421

Qy      457 NQSLDA 462
            : :::|
Db      422 hlavea 427


Search completed: June 14, 2001, 08:07:36
Job time: 188 sec
```

```
                         GenCore version 4.5
                Copyright (c) 1993 - 2000  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:          June 13, 2001, 15:39:37 ; Search time 84.7 Seconds
                                              (without alignments)
                                              4324.869 Million cell updates/sec

Title:           US-09-445-223-2
Perfect score:   2098
Sequence:        1 ggccattatggatggatggg..........aaaaaaaaaaaaaaaaaaa 2098

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        302621 seqs, 87301344 residues

Total number of hits satisfying chosen parameters:      605242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_NA:*
                1:  /cgn1_7/ptodata/1/ina/5A_COMB.seq:*
                2:  /cgn1_7/ptodata/1/ina/5B_COMB.seq:*
                3:  /cgn1_7/ptodata/1/ina/6A_COMB.seq:*
                4:  /cgn1_7/ptodata/1/ina/6B_COMB.seq:*
                5:  /cgn1_7/ptodata/1/ina/PCTUS_COMB.seq:*
                6:  /cgn1_7/ptodata/1/ina/backfiles1.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                      %
Result              Query
  No.     Score     Match Length DB  ID                       Description
 -----------------------------------------------------------------------------
     1   1865.6    88.9    1931  3  US-09-019-942-2          Sequence 2, Appli
     2      210    10.0     210  1  US-08-700-575-31         Sequence 31, Appl
     3    102.4     4.9    3516  4  US-09-188-930-257        Sequence 257, App
     4     81.4     3.9    2890  1  US-07-928-464-1          Sequence 1, Appli
     5     81.4     3.9    2890  5  PCT-US93-07347-1         Sequence 1, Appli
     6     81.4     3.9    3033  1  US-08-003-311B-1         Sequence 1, Appli
     7     81.4     3.9    3033  1  US-08-261-432-1          Sequence 1, Appli
     8     78.2     3.7    1888  4  US-09-188-930-66         Sequence 66, Appl
     9     61.8     2.9    1554  2  US-08-587-680A-24        Sequence 24, Appl
    10       58     2.8    1558  1  US-08-455-550-7          Sequence 7, Appli
    11     54.2     2.6    1557  3  US-09-329-418-2          Sequence 2, Appli
    12     54.2     2.6    1873  3  US-09-329-418-1          Sequence 1, Appli
    13     53.6     2.6    1582  3  US-08-545-196B-10        Sequence 10, Appl
    14     53.6     2.6    1582  3  US-08-545-196B-12        Sequence 12, Appl
c   15     53.2     2.5    5852  1  US-07-867-106-2          Sequence 2, Appli
c   16     52.8     2.5   19124  2  US-08-487-826B-13        Sequence 13, Appl
    17       52     2.5    1074  3  US-09-248-335-67         Sequence 67, Appl
    18       52     2.5    1364  1  US-08-265-087-3          Sequence 3, Appli
    19       52     2.5    1364  1  US-08-621-493-3          Sequence 3, Appli
    20       52     2.5    1364  2  US-08-965-688-3          Sequence 3, Appli
    21       52     2.5    1364  4  US-09-260-173-3          Sequence 3, Appli
    22     51.2     2.4    1641  1  US-08-300-903A-8         Sequence 8, Appli
    23       51     2.4    1241  1  US-07-593-657-6          Sequence 6, Appli
    24     50.4     2.4     958  2  US-08-757-046A-5         Sequence 5, Appli
    25     50.4     2.4     958  4  US-09-447-208-5          Sequence 5, Appli
    26     50.4     2.4     958  4  US-09-135-988-5          Sequence 5, Appli
    27     50.4     2.4    2852  3  US-09-027-137-2          Sequence 2, Appli
```

```
    28     49.6     2.4     991  3  US-08-924-747-25         Sequence 25, Appl
    29     49.6     2.4     991  4  US-09-247-373B-25        Sequence 25, Appl
    30     49.6     2.4     991  4  US-09-296-715-25         Sequence 25, Appl
    31     49.6     2.4   11517  1  US-07-920-281C-1         Sequence 1, Appli
    32     49.6     2.4   11517  4  US-08-466-277-1          Sequence 1, Appli
    33     49.2     2.3    5173  1  US-08-242-677-1          Sequence 1, Appli
    34     48.6     2.3   19124  2  US-08-487-826B-13        Sequence 13, Appl
    35     48.4     2.3    1117  4  US-09-247-373B-33        Sequence 33, Appl
    36     48.2     2.3    1562  1  US-08-330-108-1          Sequence 1, Appli
    37     48.2     2.3    1562  5  PCT-US92-10087-1         Sequence 1, Appli
c   38     48.2     2.3    2760  1  US-08-101-593-1          Sequence 1, Appli
    39       48     2.3    2550  6  5258287-23              Patent No. 5258287
    40       48     2.3    3207  1  US-07-946-497-1          Sequence 1, Appli
    41       48     2.3    3207  1  US-08-483-322-1          Sequence 1, Appli
    42       48     2.3    3207  2  US-08-478-882-1          Sequence 1, Appli
    43     47.8     2.3     731  1  US-08-451-405A-2         Sequence 2, Appli
    44     47.4     2.3    2296  4  US-08-496-841C-137       Sequence 137, App
    45     47.2     2.2     140  1  US-08-628-417-5          Sequence 5, Appli

                                ALIGNMENTS

RESULT   1
US-09-019-942-2
; Sequence 2, Application US/09019942
; Patent No. 6033855
;  GENERAL INFORMATION:
;    APPLICANT:  Bertin, John
;    TITLE OF INVENTION:  GENES ENCODING CASPASE RECRUITMENT
;    TITLE OF INVENTION:  DOMAIN POLYPEPTIDES
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/019,942
;      FILING DATE:  06-FEB-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meiklejohn, Ph.D., Anita L.
;      REGISTRATION NUMBER:  35,283
;      REFERENCE/DOCKET NUMBER:  07334/068001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1931 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-019-942-2

  Query Match            88.9%;  Score 1865.6;  DB 3;  Length 1931;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches 1868;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy     58 gtcagctctggttcggagaagcagcggctggcgtgggccatccggggaatgggcgccctc 117
```

```
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12 GTCAGCTCTGGTTCGGAGAAGCAGCGGCTGGCGTGGGCCATCCGGGGAATGGGCGCCCTC 71

Qy     118 gtgacctagtgttgcggggcaaaaagggtcttgccggcctcgctcgtgcaggggcgtatc 177
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72 GTGACCTAGTGTTGCGGGGCAAAAAGGGTCTTGCCGGCCTCGCTCGTGCAGGGGCGTATC 131

Qy     178 tgggcgcctgagcgcggcgtgggagccttgggagccgccgcagcaggggcacacccgga 237
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     132 TGGGCGCCTGAGCGCGGCGTGGGAGCCTTGGGAGCCGCCGCAGCAGGGGCACACCCGGA 191

Qy     238 accggcctgagcgcccgggaccatgaacggggaggccatctgcagcgccctgcccaccat 297
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     192 ACCGGCCTGAGCGCCCGGGACCATGAACGGGGAGGCCATCTGCAGCGCCCTGCCCACCAT 251

Qy     298 tccctaccacaaactcgccgacctgcgctacctgagccgcggcgcctctggcactgtgtc 357
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     252 TCCCTACCACAAACTCGCCGACCTGCGCTACCTGAGCCGCGGCGCCTCTGGCACTGTGTC 311

Qy     358 gtccgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctgcacatccacac 417
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     312 GTCCGCCCGCCACGCAGACTGGCGCGTCCAGGTGGCCGTGAAGCACCTGCACATCCACAC 371

Qy     418 tccgctgctcgacagtgaaagaaaggatgttttaagagaagctgaaattttacacaaagc 477
           |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db     372 TCCGCTGCTCGACAGTGAAAGAAAGGATGTCTTAAGAGAAGCTGAAATTTTACACAAAGC 431

Qy     478 tagatttagttacatttttccaattttgggaatttgcaatgagcctgaatttttgggaat 537
           |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db     432 TAGATTTAGTTACATTCTTCCAATTTTGGGAATTTGCAATGAGCCTGAATTTTTGGGAAT 491

Qy     538 agttactgaatacatgccaaatggatcattaaatgaactcctacataggaaaactgaata 597
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     492 AGTTACTGAATACATGCCAAATGGATCATTAAATGAACTCCTACATAGGAAAACTGAATA 551

Qy     598 tcctgatgttgcttggccattgagatttcgcatcctgcatgaaattgcccttggtgtaaa 657
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     552 TCCTGATGTTGCTTGGCCATTGAGATTTCGCATCCTGCATGAAATTGCCCTTGGTGTAAA 611

Qy     658 ttacctgcacaatatgactcctcctttacttcatcatgacttgaagactcagaatatctt 717
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     612 TTACCTGCACAATATGACTCCTCCTTTACTTCATCATGACTTGAAGACTCAGAATATCTT 671

Qy     718 attggacaatgaatttcatgttaagattgcagattttggtttatcaaagtggcgcatgat 777
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     672 ATTGGACAATGAATTTCATGTTAAGATTGCAGATTTTGGTTTATCAAAGTGGCGCATGAT 731

Qy     778 gtccctctcacagtcacgaagtagcaaatctgcaccagaaggagggacaattatttatat 837
           |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db     732 GTCCCTCTCACAGTCACGAAGTAGCAAATCTGCACCAGAAGGAGGGACAATTATCTATAT 791

Qy     838 gccacctgaaaactatgaacctggacaaaaatcaagggccagtatcaagcacgatatata 897
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     792 GCCACCTGAAAACTATGAACCTGGACAAAAATCAAGGGCCAGTATCAAGCACGATATATA 851

Qy     898 tagctatgcagttatcacatgggaagtgttatccagaaaacagccttttgaagatgtcac 957
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     852 TAGCTATGCAGTTATCACATGGGAAGTGTTATCCAGAAAACAGCCTTTTGAAGATGTCAC 911

Qy     958 caatcctttgcagataatgtatagtgtgtcacaaggacatcgacctgttattaatgaaga 1017
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     912 CAATCCTTTGCAGATAATGTATAGTGTGTCACAAGGACATCGACCTGTTATTAATGAAGA 971

Qy    1018 aagtttgccatatgatatacctcaccgagcacgtatgatctctctaatagaaagtggatg 1077
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     972 AAGTTTGCCATATGATATACCTCACCGAGCACGTATGATCTCTCTAATAGAAAGTGGATG 1031

Qy    1078 ggcacaaaatccagatgaaagaccatctttcttaaaatgtttaatagaacttgaaccagt 1137
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1032 GGCACAAAATCCAGATGAAAGACCATCTTTCTTAAAATGTTTAATAGAACTTGAACCAGT 1091

Qy    1138 tttgagaacatttgaagagataacttttcttgaagctgttattcagctaaagaaaacaaa 1197
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1092 TTTGAGAACATTTGAAGAGATAACTTTTCTTGAAGCTGTTATTCAGCTAAAGAAAACAAA 1151

Qy    1198 gttacagagtgtttcaagtgccattcacctatgtgacaagaagaaaatggaattatctct 1257
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1152 GTTACAGAGTGTTTCAAGTGCCATTCACCTATGTGACAAGAAGAAAATGGAATTATCTCT 1211

Qy    1258 gaacatacctgtaaatcatggtccacaagaggaatcatgtggatcctctcagctccatga 1317
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1212 GAACATACCTGTAAATCATGGTCCACAAGAGGAATCATGTGGATCCTCTCAGCTCCATGA 1271

Qy    1318 aaatagtggttctcctgaaacttcaaggtccctgccagctcctcaagacaatgatttttt 1377
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1272 AAATAGTGGTTCTCCTGAAACTTCAAGGTCCCTGCCAGCTCCTCAAGACAATGATTTTTT 1331

Qy    1378 atctagaaaagctcaagactgttattttatgaagctgcatcactgtcctggaaatcacag 1437
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1332 ATCTAGAAAAGCTCAAGACTGTTATTTTATGAAGCTGCATCACTGTCCTGGAAATCACAG 1391

Qy    1438 ttgggatagcaccatttctggatctcaaagggctgcattctgtgatcacaagaccactcc 1497
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    1392 TTGGGATAGCACCATTTCTGGATCTCAAAGGGCTGCATTCTGTGATCACAAGACCATTCC 1451

Qy    1498 atgctcttcagcaataataaatccactctcaactgcaggaaactcagaacgtctgcagcc 1557
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1452 ATGCTCTTCAGCAATAATAAATCCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAGCC 1511

Qy    1558 tggtatagcccagcagtggatccagagcaaaagggaagacattgtgaaccaaatgacaga 1617
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1512 TGGTATAGCCCAGCAGTGGATCCAGAGCAAAAGGGAAGACATTGTGAACCAAATGACAGA 1571

Qy    1618 agcctgccttaaccagtcgctagatgcccttctgtccagggacttgatcatgaaagagga 1677
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1572 AGCCTGCCTTAACCAGTCGCTAGATGCCCTTCTGTCCAGGGACTTGATCATGAAAGAGGA 1631

Qy    1678 ctatgaacttgttagtaccaagcctacaaggacctcaaaagtcagacaattactagacac 1737
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1632 CTATGAACTTGTTAGTACCAAGCCTACAAGGACCTCAAAAGTCAGACAATTACTAGACAC 1691

Qy    1738 tactgacatccaaggagaagaatttgccaaagttatagtacaaaaattgaaagataacaa 1797
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1692 TACTGACATCCAAGGAGAAGAATTTGCCAAAGTTATAGTACAAAAATTGAAAGATAACAA 1751

Qy    1798 acaaatgggtcttcagccttacccggaaatacttgtggtttctagatcaccatctttaaa 1857
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1752 ACAAATGGGTCTTCAGCCTTACCCGGAAATACTTGTGGTTTCTAGATCACCATCTTTAAA 1811

Qy    1858 tttacttcaaaataaaagcatgtaagtgactgtttttcaagaagaaatgtgtttcataaa 1917
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1812 TTTACTTCAAAATAAAAGCATGTAAGTGACTGTTTTTCAAGAAGAAATGTGTTTCATAAA 1871

Qy    1918 aggatatttata 1929
           ||||||||||||
Db    1872 AGGATATTTATA 1883

RESULT   2
US-08-700-575-31
; Sequence 31, Application US/08700575
; Patent No. 5817479
;  GENERAL INFORMATION:
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Hawkins, Phillip R.
;    APPLICANT:  Wilde, Craig G.
;    TITLE OF INVENTION:  NOVEL HUMAN KINASE HOMOLOGS
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/700,575
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BILLINGS, LUCY J
;      REGISTRATION NUMBER:  36749
;      REFERENCE/DOCKET NUMBER:  SP-100 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  210 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      LIBRARY:  THP-1 Phorbol LPS
;      CLONE:  156108
US-08-700-575-31


  Query Match             10.0%;  Score 210;  DB 1;  Length 210;
  Best Local Similarity   100.0%;  Pred. No. 4.4e-39;
  Matches  210;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    844 tgaaaactatgaacctggacaaaaatcaagggccagtatcaagcacgatatatatagcta 903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TGAAAACTATGAACCTGGACAAAAATCAAGGGCCAGTATCAAGCACGATATATATAGCTA 60

Qy    904 tgcagttatcacatgggaagtgttatccagaaaacagccttttgaagatgtcaccaatcc 963
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TGCAGTTATCACATGGGAAGTGTTATCCAGAAAACAGCCTTTTGAAGATGTCACCAATCC 120

Qy    964 tttgcagataatgtatagtgtgtcacaaggacatcgacctgttattaatgaagaaagttt 1023
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TTTGCAGATAATGTATAGTGTGTCACAAGGACATCGACCTGTTATTAATGAAGAAAGTTT 180

Qy   1024 gccatatgatatacctcaccgagcacgtat 1053
          ||||||||||||||||||||||||||||||
Db    181 GCCATATGATATACCTCACCGAGCACGTAT 210

RESULT   3
US-09-188-930-257
; Sequence 257, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
;  APPLICANT: Watson, James D.
;  APPLICANT: Strachan, Lorna
;  APPLICANT: Sleeman, Matthew
;  APPLICANT: Onrust, Rene
;  APPLICANT: Murison, James Greg
;  TITLE OF INVENTION: Compositions Isolated From Skin Cells
;  TITLE OF INVENTION: and Methods For Their Use
;  FILE REFERENCE: 11000.1011c1
;  CURRENT APPLICATION NUMBER: US/09/188,930A
;  CURRENT FILING DATE: 1998-11-09
;  NUMBER OF SEQ ID NOS: 348
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 257
;   LENGTH: 3516
;   TYPE: DNA
;   ORGANISM: Mouse
US-09-188-930-257
```

```
  Query Match             4.9%;  Score 102.4;  DB 4;  Length 3516;
  Best Local Similarity   51.8%;  Pred. No. 3.4e-14;
  Matches  372;  Conservative    0;  Mismatches  316;  Indels   30;  Gaps    5;

Qy    287 ctgcccaccattccctaccacaaactcgccgacctgcgctacctgagccgcggcgcctct 346
          |||| |||| |   |  |  | || |||| | |  |    |  || ||   |||| ||
Db     43 ctgcgcaccttcgacgccggcgaattcgcaggctgggagaaggtgggctcgggcggcttc 102

Qy    347 ggcactgtgtcgtccgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctg 406
          ||    ||||     |  ||||| |   |||||        | | ||  | |||  |  |
Db    103 gggcaggtgtacaaggtgcgccatgtgcactggaagacgtggctcgcgatcaagtgctcg 162

Qy    407 cacatccacactccgctgctcgacagtgaaagaaaggatgttttaagagaagctgaaatt 466
          | ||  |     |   |   |||||| ||| ||| |||  |  |    |||||| | |
Db    163 cccagtc---tgcacgtcgacgacagggaacgaatggagctcctggaggaagctaagaag 219

Qy    467 ttacacaaagctagatttagttacatttttccaattttgggaatttgcaatgagcctgaa 526
           |  | |  || |  ||  | |||||| | ||  | |  || || ||| | || ||||
Db    220 atggagatggccaagttccgatacattctacctgtgtacggcatatgccaggaacctg-- 277

Qy    527 tttttgggaatagttactgaatacatgccaaatggatcattaaatgaactcctacatagg 586
              | ||  | || |  || ||||||   |  || ||  |  |  | || ||
Db    278 ----tcggcttggtcatggagtacatggagacaggctccctggagaagctgctggcctca 333

Qy    587 aaaactgaatatcctgatgttgcttggccattgagatttcgcatcctgcatgaaattgcc 646
           |  |              |  |||||    || | ||||||||| |||| || |  |||
Db    334 gagccat------------tgccttgggacctgcgctttcgcatcgtgcacgagacagcc 381

Qy    647 cttggtgtaaattacctgcacaatatgactcctcctttacttcatcatgacttgaagact 706
           | ||  | || | ||||||    ||| |||| ||  | || | || |  ||| ||||| |
Db    382 gtgggcatgaacttcctgcattgcatgtctccgccactgctgcacctagacctgaagcca 441

Qy    707 cagaatatcttattggacaatgaatttcatgttaagattgcagattttggtttatcaaag 766
            ||| ||| |  ||||      | | ||||| |||||| | || |||||  |  | |||
Db    442 gcgaacatcctgctggatgcccactaccatgtcaagatttctgactttgggctggccaag 501

Qy    767 tggcgcatgatgtccctctcacagtcacgaagtagcaaatctgcaccagaaggagggaca 826
          ||       ||||||| ||| ||    |  || |   |                || |||
Db    502 tgcaatggcatgtcccactctcatgacctcagcatggatggcctgttt------ggtaca 555

Qy    827 attatttatatgccacctgaaaactatgaacctggacaaaaatcaagggccagtatcaag 886
          ||   |||  | || || ||      |   | ||   | |       |       | |||
Db    556 atcgcttacctccctccagagcgaatt---cgtgagaagagccgcttgtttgacaccaaa 612

Qy    887 cacgatatatatagctatgcagttatcacatgggaagtgttatccagaaaacagcctttt 946
          || ||| |||| ||||  ||  || | |  ||||  ||| |  |    ||  |||| |||
Db    613 catgatgtatacagcttcgccattgtgatctggggtgtgcttacacagaagaagccattt 672

Qy    947 gaagatgtcaccaatcctttgcagataatgtatagtgtgtcacaaggacatcgacctg 1004
          | |||||  |  ||     | || || |||   |  |||  | | || || || || |
Db    673 gcagatgaaaagaacatcctacacatcatgatgaaagtggtaaagggccaccgcccag 730

RESULT   4
US-07-928-464-1
; Sequence 1, Application US/07928464
; Patent No. 5367065
;  GENERAL INFORMATION:
;    APPLICANT:  Ecker, Joseph R.
;    APPLICANT:  Kieber, Joseph J.
;    TITLE OF INVENTION:  Constitutive Triple Response Gene and
;    TITLE OF INVENTION:  Mutations
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock, Washburn, Kurtz, Mackiewicz and
;      ADDRESSEE:  No. 5367065ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  U.S.A.
;      ZIP:  19103
;    COMPUTER READABLE FORM:
```

```
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/928,464
;      FILING DATE:  19920810
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Miller, Suzanne E.
;      REGISTRATION NUMBER:  32,279
;      REFERENCE/DOCKET NUMBER:  UPN-1086
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2890 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  118..2583
US-07-928-464-1

  Query Match             3.9%;  Score 81.4;  DB 1;  Length 2890;
  Best Local Similarity   50.1%;  Pred. No. 1.8e-09;
  Matches  202;  Conservative    0;  Mismatches  201;  Indels    0;  Gaps    0;

Qy    360 ccgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctgcacatccacactc 419
          | | || ||  || || ||||  | |  ||  |  |    | |  | | |||  |
Db   1808 CTGTCCACCGTGCTGAGTGGCATGGCTCGGATGTTGCTGTGAAAATTCTCATGGAGCAAG 1867

Qy    420 cgctgctcgacagtgaaagaaaggatgttttaagagaagctgaaattttacacaaagcta 479
             | |  |           || ||  | |||||||| | ||  ||  |  |      |
Db   1868 ACTTCCATGCTGAGCGTGTTAATGAGTTCTTAAGAGAGGTTGCGATAATGAAACGCCTTC 1927

Qy    480 gatttagttacatttttccaattttgggaatttgcaatgagcctgaatttttgggaatag 539
          |      | ||||| |||   |  ||||      || | | |||  |  ||||  |||||
Db   1928 GCCACCCTAACATTGTTCTCTTCATGGGTGCGGTCACTCAACCTCCAAATTTGTCAATAG 1987

Qy    540 ttactgaatacatgccaaatggatcattaaatgaactcctacataggaaaactgaatatc 599
          | || |||||  || |||  ||    ||| |   |||  | ||||  |     | |
Db   1988 TGACAGAATATTTGTCAAGAGGTAGTTTATACAGACTTTTGCATAAAAGTGGAGCAAGGG 2047

Qy    600 ctgatgttgcttggccattgagatttcgcatcctgcatgaaattgcccttggtgtaaatt 659
             |  | | |  |  |    |  |  | ||     ||||  | ||    ||  | ||||
Db   2048 AGCAATTAGATGAGAGACGTCGCCTGAGTATGGCTTATGATGTGGCTAAGGGAATGAATT 2107

Qy    660 acctgcacaatatgactcctcctttacttcatcatgacttgaagactcagaatatcttat 719
          | || ||||||   | ||||||  |  | |||   ||  | ||  |||  ||  | ||
Db   2108 ATCTTCACAATCGCAATCCTCCAATTGTGCATAGAGATCTAAAATCTCCAAACTTATTGG 2167

Qy    720 tggacaatgaatttcatgttaagattgcagattttggtttatc 762
          | |||||  ||| |   || ||| ||   ||||||||| | ||
Db   2168 TTGACAAAAAATATACAGTCAAGGTTTGTGATTTTGGTCTCTC 2210

RESULT   5
PCT-US93-07347-1
; Sequence 1, Application PC/TUS9307347
;  GENERAL INFORMATION:
;    APPLICANT:  Ecker, Joseph R.
;    APPLICANT:  Kieber, Joseph J.
;    TITLE OF INVENTION:  Constitutive Triple Response Gene and
;    TITLE OF INVENTION:  Mutations
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock, Washburn, Kurtz, Mackiewicz and
;      ADDRESSEE:  Norris
;      STREET:  One Liberty Place - 46th Floor
```

```
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  U.S.A.
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/07347
;      FILING DATE:  19930805
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Miller, Suzanne E.
;      REGISTRATION NUMBER:  32,279
;      REFERENCE/DOCKET NUMBER:  UPN-1086
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2890 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  118..2583
PCT-US93-07347-1

  Query Match             3.9%;  Score 81.4;  DB 5;  Length 2890;
  Best Local Similarity   50.1%;  Pred. No. 1.8e-09;
  Matches  202;  Conservative    0;  Mismatches  201;  Indels    0;  Gaps    0;

Qy    360 ccgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctgcacatccacactc 419
          | | || ||  || || ||||  | |  ||  |  |    | |  | | |||  |
Db   1808 CTGTCCACCGTGCTGAGTGGCATGGCTCGGATGTTGCTGTGAAAATTCTCATGGAGCAAG 1867

Qy    420 cgctgctcgacagtgaaagaaaggatgttttaagagaagctgaaattttacacaaagcta 479
             | |  |           || ||  | |||||||| | ||  ||  |  |      |
Db   1868 ACTTCCATGCTGAGCGTGTTAATGAGTTCTTAAGAGAGGTTGCGATAATGAAACGCCTTC 1927

Qy    480 gatttagttacatttttccaattttgggaatttgcaatgagcctgaatttttgggaatag 539
          |      | ||||| |||   |  ||||      || | | |||  |  ||||  |||||
Db   1928 GCCACCCTAACATTGTTCTCTTCATGGGTGCGGTCACTCAACCTCCAAATTTGTCAATAG 1987

Qy    540 ttactgaatacatgccaaatggatcattaaatgaactcctacataggaaaactgaatatc 599
          | || |||||  || |||  ||    ||| |   |||  | ||||  |     | |
Db   1988 TGACAGAATATTTGTCAAGAGGTAGTTTATACAGACTTTTGCATAAAAGTGGAGCAAGGG 2047

Qy    600 ctgatgttgcttggccattgagatttcgcatcctgcatgaaattgcccttggtgtaaatt 659
             |  | | |  |  |    |  |  | ||     ||||  | ||    ||  | ||||
Db   2048 AGCAATTAGATGAGAGACGTCGCCTGAGTATGGCTTATGATGTGGCTAAGGGAATGAATT 2107

Qy    660 acctgcacaatatgactcctcctttacttcatcatgacttgaagactcagaatatcttat 719
          | || ||||||   | ||||||  |  | |||   ||  | ||  |||  ||  | ||
Db   2108 ATCTTCACAATCGCAATCCTCCAATTGTGCATAGAGATCTAAAATCTCCAAACTTATTGG 2167

Qy    720 tggacaatgaatttcatgttaagattgcagattttggtttatc 762
          | |||||  ||| |   || ||| ||   ||||||||| | ||
Db   2168 TTGACAAAAAATATACAGTCAAGGTTTGTGATTTTGGTCTCTC 2210

RESULT   6
US-08-003-311B-1
; Sequence 1, Application US/08003311B
; Patent No. 5444166
;  GENERAL INFORMATION:
;    APPLICANT:  Ecker, Joseph R.
;    APPLICANT:  Kieber, Joseph J.
;    TITLE OF INVENTION:  Constitutive Triple Response Gene
```

```
;    TITLE OF INVENTION:  and Mutations
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz and
;      ADDRESSEE:  No. 5444166ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  U.S.A.
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/003,311B
;      FILING DATE:  January 12, 1993
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/928,464
;      FILING DATE:  August 10, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lori Y. Beardell
;      REGISTRATION NUMBER:  34,293
;      REFERENCE/DOCKET NUMBER:  UPN-1108
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3033 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  118..2583
US-08-003-311B-1

  Query Match            3.9%;  Score 81.4;  DB 1;  Length 3033;
  Best Local Similarity   50.1%;  Pred. No. 1.9e-09;
  Matches  202;  Conservative    0;  Mismatches  201;  Indels    0;  Gaps    0;

Qy    360 ccgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctgcacatccacactc 419
          | | || ||  || || ||||  | |  ||  |  |    | |  | | |||  |
Db   1808 CTGTCCACCGTGCTGAGTGGCATGGCTCGGATGTTGCTGTGAAAATTCTCATGGAGCAAG 1867

Qy    420 cgctgctcgacagtgaaagaaaggatgttttaagagaagctgaaattttacacaaagcta 479
             | |  |           || ||  | |||||||| | ||  ||  |  |      |
Db   1868 ACTTCCATGCTGAGCGTGTTAATGAGTTCTTAAGAGAGGTTGCGATAATGAAACGCCTTC 1927

Qy    480 gatttagttacatttttccaattttgggaatttgcaatgagcctgaatttttgggaatag 539
          |      | ||||| |||   |  ||||      || | | |||  |  ||||  |||||
Db   1928 GCCACCCTAACATTGTTCTCTTCATGGGTGCGGTCACTCAACCTCCAAATTTGTCAATAG 1987

Qy    540 ttactgaatacatgccaaatggatcattaaatgaactcctacataggaaaactgaatatc 599
          | || |||||  || |||  ||    ||| |   |||  | ||||  |     | |
Db   1988 TGACAGAATATTTGTCAAGAGGTAGTTTATACAGACTTTTGCATAAAAGTGGAGCAAGGG 2047

Qy    600 ctgatgttgcttggccattgagatttcgcatcctgcatgaaattgcccttggtgtaaatt 659
             |  | | |  |  |    |  |  | ||     ||||  | ||    ||  | ||||
Db   2048 AGCAATTAGATGAGAGACGTCGCCTGAGTATGGCTTATGATGTGGCTAAGGGAATGAATT 2107

Qy    660 acctgcacaatatgactcctcctttacttcatcatgacttgaagactcagaatatcttat 719
          | || ||||||   | ||||||  |  | |||   ||  | ||  |||  ||  | ||
Db   2108 ATCTTCACAATCGCAATCCTCCAATTGTGCATAGAGATCTAAAATCTCCAAACTTATTGG 2167

Qy    720 tggacaatgaatttcatgttaagattgcagattttggtttatc 762
          | |||||  ||| |   || ||| ||   ||||||||| | ||
Db   2168 TTGACAAAAAATATACAGTCAAGGTTTGTGATTTTGGTCTCTC 2210
```

```
RESULT   7
US-08-261-432-1
; Sequence 1, Application US/08261432
; Patent No. 5602322
;  GENERAL INFORMATION:
;    APPLICANT:  Ecker, Joseph R.
;    APPLICANT:  Kieber, Joseph J.
;    TITLE OF INVENTION:  Constitutive Triple Response Gene
;    TITLE OF INVENTION:  and Mutations
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz and
;      ADDRESSEE:  No. 5602322ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  U.S.A.
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/261,432
;      FILING DATE:  June 17, 1994
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/003,311
;      FILING DATE:  January 12, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lori Y. Beardell
;      REGISTRATION NUMBER:  34,293
;      REFERENCE/DOCKET NUMBER:  UPN-1864
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3033 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  118..2583
US-08-261-432-1

  Query Match            3.9%;  Score 81.4;  DB 1;  Length 3033;
  Best Local Similarity   50.1%;  Pred. No. 1.9e-09;
  Matches  202;  Conservative    0;  Mismatches  201;  Indels    0;  Gaps    0;

Qy    360 ccgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctgcacatccacactc 419
          | | || ||  || || ||||  | |  ||  |  |    | |  | | |||  |
Db   1808 CTGTCCACCGTGCTGAGTGGCATGGCTCGGATGTTGCTGTGAAAATTCTCATGGAGCAAG 1867

Qy    420 cgctgctcgacagtgaaagaaaggatgttttaagagaagctgaaattttacacaaagcta 479
             | |  |           || ||  | |||||||| | ||  ||  |  |      |
Db   1868 ACTTCCATGCTGAGCGTGTTAATGAGTTCTTAAGAGAGGTTGCGATAATGAAACGCCTTC 1927

Qy    480 gatttagttacatttttccaattttgggaatttgcaatgagcctgaatttttgggaatag 539
          |      | ||||| |||   |  ||||      || | | |||  |  ||||  |||||
Db   1928 GCCACCCTAACATTGTTCTCTTCATGGGTGCGGTCACTCAACCTCCAAATTTGTCAATAG 1987

Qy    540 ttactgaatacatgccaaatggatcattaaatgaactcctacataggaaaactgaatatc 599
          | || |||||  || |||  ||    ||| |   |||  | ||||  |     | |
Db   1988 TGACAGAATATTTGTCAAGAGGTAGTTTATACAGACTTTTGCATAAAAGTGGAGCAAGGG 2047

Qy    600 ctgatgttgcttggccattgagatttcgcatcctgcatgaaattgcccttggtgtaaatt 659
```

```
              |  | | |  |  |    |  |  | ||     ||||  | ||    ||  | ||||
Db    2048 AGCAATTAGATGAGAGACGTCGCCTGAGTATGGCTTATGATGTGGCTAAGGGAATGAATT 2107

Qy     660 acctgcacaatatgactcctcctttacttcatcatgacttgaagactcagaatatcttat 719
           | || ||||||   | ||||||  |  | |||   ||  | ||  |||  ||  | ||
Db    2108 ATCTTCACAATCGCAATCCTCCAATTGTGCATAGAGATCTAAAATCTCCAAACTTATTGG 2167

Qy     720 tggacaatgaatttcatgttaagattgcagattttggtttatc 762
           | |||||  ||| |   || ||| ||   ||||||||| | ||
Db    2168 TTGACAAAAAATATACAGTCAAGGTTTGTGATTTTGGTCTCTC 2210


RESULT   8
US-09-188-930-66
; Sequence 66, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
;  APPLICANT: Watson, James D.
;  APPLICANT: Strachan, Lorna
;  APPLICANT: Sleeman, Matthew
;  APPLICANT: Onrust, Rene
;  APPLICANT: Murison, James Greg
;  TITLE OF INVENTION: Compositions Isolated From Skin Cells
;  TITLE OF INVENTION: and Methods For Their Use
;  FILE REFERENCE: 11000.1011c1
;  CURRENT APPLICATION NUMBER: US/09/188,930A
;  CURRENT FILING DATE: 1998-11-09
;  NUMBER OF SEQ ID NOS: 348
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 66
;   LENGTH: 1888
;   TYPE: DNA
;   ORGANISM: mouse
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1690)...(1690)
;   NAME/KEY: unsure
;   LOCATION: (1755)...(1755)
;   NAME/KEY: unsure
;   LOCATION: (1864)...(1864)
US-09-188-930-66


  Query Match             3.7%;  Score 78.2;  DB 4;  Length 1888;
  Best Local Similarity   50.6%;  Pred. No. 8.1e-09;
  Matches  365;  Conservative    0;  Mismatches  323;  Indels   33;  Gaps    6;

Qy     287 ctgcccaccattccctaccacaaactcgccgacctgcgctacctgagccgcggcgcctct 346
           |||| |||| |   |  |  | || |||| | |  |    |  || ||   |||| ||
Db      43 ctgcgcaccttcgacgccggcgaattcgcaggctgggagaaggtgggctcgggcggcttc 102

Qy     347 ggcactgtgtcgtccgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctg 406
           ||    ||||     |  ||||| |   |||||        | | ||  | |||  |  |
Db     103 gggcaggtgtacaaggtgcgccatgtgcactggaagacgtggctcgcgatcaagtgctcg 162

Qy     407 cacatccacactccgctgctcgacagtgaaagaaaggatgttttaagagaagctgaaatt 466
           | ||  |     |   |   |||||| ||| ||| |||  |  |    |||||| | |
Db     163 cccagtc---tgcacgtcgacgacagggaacgaatggagctcctggaggaagctaagaag 219

Qy     467 ttacacaaagctagatttagttacattttttccaattttgggaatttgcaatgagcctgaa 526
            |  | |  || |  ||  | |||||| | ||  | |  || || ||| | || ||||
Db     220 atggagatggccaagttccgatacattctacctgtgtacggcatatgccaggaacctg-- 277

Qy     527 tttttgggaatagttactgaatacatgccaaatggatcattaaatgaactcctacatagg 586
               | ||  | || |  || ||||||   |  || ||  |  |  | || ||
Db     278 ----tcggcttggtcatggagtacatggagacaggctccctggagaagctgctggcctca 333

Qy     587 aaaactgaatatcctgatgttgcttggccattgagatttcgcatcctgcatgaaattgcc 646
            |  |              |  |||||    || | ||||||||| |||| || |  |||
Db     334 gagccat------------tgccttgggacctgcgctttcgcatcgtgcacgagacagcc 381

Qy     647 cttggtgtaaattacctgcacaatatgactcctcctttacttcatcatgacttgaagact 706
            | ||  | || | ||||||    ||| |||| ||  | || || |  ||| ||||| |
Db     382 gtgggcatgaacttcctgcattgcatgtctccgccactgctgcacctagacctgaagcca 441

Qy     707 cagaatatcttattggacaatgaatttcatgt---taagattgcagattttggtttatca 763
             ||| |||||  ||||     | |  ||  |    |   ||   || |||||  |  |
Db     442 gcgaacatcttgctggatgcccactaccaaatgtcaagatttcttgactttgggctggcc 501

Qy     764 aagtggcgcatgatgtccctctcacagtcacgaagtagcaaatctgcaccagaaggaggg 823
           |||||       ||||||| ||| ||    |  || |        |           ||
Db     502 aagtgcaatggcatgtcccactctcatgacctcagcat------ggatggcctgtttggt 555

Qy     824 acaattatttatatgccacctgaaaactatgaacctggacaaaaatcaagggccagtatc 883
           |||||    ||  | || || ||      |   | ||   | |       |      | |
Db     556 acaatcggctacctccctccagagcgaatt---cgtgagaagagccgcttgtttgacacc 612

Qy     884 aagcacgatatatatagctatgcagttatcacatgggaagtgttatccagaaaacagcct 943
           || || ||| |||| ||||  ||  || | |  ||||  ||| |  |    ||  | ||
Db     613 aaacatgatgtatacagcttcgccattgtgatctggggtgtgcttacacagaataatcca 672

Qy     944 tttgaagatgtcaccaatcctttgcagataatgtatagtgtgtcacaaggacatcgacct 1003
           |||| |||||  |  ||     | || || |||   |  |||  | | || || || ||
Db     673 tttgcagatgaaaagaacatcctacacatcatgatgaaagtggtaaagggccaccgccca 732

Qy    1004 g 1004
           |
Db     733 g 733


RESULT   9
US-08-587-680A-24
; Sequence 24, Application US/08587680A
; Patent No. 5977434
;  GENERAL INFORMATION:
;    APPLICANT:  Ronald, Pamela C.
;    APPLICANT:  Wang, Guo-Liang
;    APPLICANT:  Song, Wen-Yuang
;    APPLICANT:  Szabo, Veronique
;    TITLE OF INVENTION:  Procedures and Materials for Conferring
;    TITLE OF INVENTION:  Disease Resistance in Plants
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/587,680A
;      FILING DATE:  17-JAN-1996
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/373,375
;      FILING DATE:  17-JAN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/475,891
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/004,645
;      FILING DATE:  29-SEP-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/567,375
;      FILING DATE:  04-DEC-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bastian, Kevin L.
;      REGISTRATION NUMBER:  34,774
;      REFERENCE/DOCKET NUMBER:  02307O-058940US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
```

```
;       TELEFAX:  (415) 576-0300
;   INFORMATION FOR SEQ ID NO:  24:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1554 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA (partial)
US-08-587-680A-24


  Query Match             2.9%;  Score 61.8;  DB 2;  Length 1554;
  Best Local Similarity   49.8%;  Pred. No. 3.9e-05;
  Matches  156;  Conservative    0;  Mismatches  157;  Indels    0;  Gaps    0;

Qy    455 gaagctgaaattttacacaaagctagatttagttacatttttccaattttgggaatttgc 514
          |||  | | |  |||   ||    ||   ||| |||||| |   | |  | |   ||||
Db    835 GAACTTAACACATTAGGGAAGATCAGGCATAGGTACATTGTGAGACTGCTCGCGTTTTGT 894

Qy    515 aatgagcctgaatttttgggaatagttactgaatacatgccaaatggatcattaaatgaa 574
              |    ||| |        |||||  ||| ||||||| |||||||   |||  ||||
Db    895 TCAAACAAGGAAGTCAACTTGCTAGTTTATGAGTACATGCTAAATGGAAGCTTAGGTGAA 954

Qy    575 ctcctacataggaaaactgaatatcctgatgttgcttggccattgagatttcgcatcctg 634
           | || ||| |||| |  |     |               |   |  |        | |
Db    955 GTGCTTCATGGGAAGAACGGCGGGCAACTCCAATGGGAAACTAGGCTAAAAATAGCCATA 1014

Qy    635 catgaaattgcccttggtgtaaattacctgcacaatatgactcctcctttacttcatcat 694
           | |              |    |||  |  ||| ||    | |||    || | || |
Db   1015 GAAGCTGCCAAGGGCCTTTCTTATTTGCACCACGATTGCTCCCCTATGATAATCCACCGC 1074

Qy    695 gacttgaagactcagaatatcttattggacaatgaatttcatgttaagattgcagatttt 754
          ||  | ||| |  | ||||| || ||| ||  |||| || | | | |  |||||||||||
Db   1075 GATGTCAAGTCCAACAATATATTGTTGAACTCTGAACTTGAAGCTCATGTTGCAGATTTT 1134

Qy    755 ggtttatcaaagt 767
          || ||| | ||||
Db   1135 GGATTAGCCAAGT 1147


RESULT  10
US-08-455-550-7
; Sequence 7, Application US/08455550
; Patent No. 5670338
;  GENERAL INFORMATION:
;    APPLICANT:  MURAKAMI, KAZUO
;    APPLICANT:  UENO, NAOTO
;    APPLICANT:  KATO, YUKIO
;    TITLE OF INVENTION:  XENOPUS LAEVIS BONE MORPHOGENETIC PROTEINS AND USE THE
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dike, Bronstein, Roberts & Cushman
;      STREET:  130 Water Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,550
;      FILING DATE:  31-MAY-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/056,564
;      FILING DATE:  30-APR-1993
;      APPLICATION NUMBER:  07/577,892
;      FILING DATE:  05-SEP-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Eisenstein, Ronald I
```

```
;      REGISTRATION NUMBER:  30628
;      REFERENCE/DOCKET NUMBER:  40302-FWC-DIV
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;      TELEX:  200291
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1558 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
US-08-455-550-7


  Query Match             2.8%;  Score 58;  DB 1;  Length 1558;
  Best Local Similarity   48.1%;  Pred. No. 0.00029;
  Matches  151;  Conservative    4;  Mismatches  159;  Indels    0;  Gaps    0;

Qy   1782 aattgaaagataacaaacaaatgggtcttcagccttacccggaaatacttgtggtttcta 1841
          | ||| | ||  |  |  || | |  |||  |    |||| ||| ||  | ||||
Db   1237 ATTTGGATGAATATGACAAAGTCGTCCTTAAAAACTACCAGGAGATGGTGGTGGAAGGGT 1296

Qy   1842 gatcaccatctttaaatttacttcaaaataaaagcatgtaagtgactgtttttcaagaag 1901
          |         || |   | |  || ||| |||||  ||| |  | |||    |||
Db   1297 GTGGGTGCCGTTGAGTCTGAGATCCAAACAAAAGACTGTTAACGGCTGGACTTCTTTCCA 1356

Qy   1902 aaatgtgtttcataaaaggatatttatatctctgttgctttgactttttttatataaaat 1961
                 ||       |   |||||||  || |  ||  |  |  ||||||  | || ||
Db   1357 CTGAACATTCACCTTGACCTTATTTATGACTTTTATGTGTAAATGTTTTTTTGACAATAT 1416

Qy   1962 ccgtgagtattaaagcttwawwraargktctttsrktaaatattagtctccctccatgac 2021
                 ||||    |  :|:: |    |   |  :||    |             |  |
Db   1417 GATCATATATTTTGACAAAATATATTTATAACTACGTATTAAAAGAAAAAAAAAAAATAA 1476

Qy   2022 actgcagtattttttttaattaatacaagtaaaaagttgaatttgaaaaaaaaaaaaaaa 2081
          | |     ||| |||| ||   | | ||  |||||    ||    |||||||||||||||
Db   1477 AATAAGTCATTATTTTAAACATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1536

Qy   2082 aaaaaaaaaaaaaa 2095
          ||||||||||||||
Db   1537 AAAAAAAAAAAAAA 1550


RESULT  11
US-09-329-418-2
; Sequence 2, Application US/09329418
; Patent No. 6096539
; GENERAL INFORMATION:
;  APPLICANT: ZENECA Limited
;  TITLE OF INVENTION: PROTEIN ACTIVATOR OF APOPTOSIS
;  FILE REFERENCE: PHM.70536
;  CURRENT APPLICATION NUMBER: US/09/329,418
;  CURRENT FILING DATE: 1999-06-11
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 1557
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-09-329-418-2


  Query Match             2.6%;  Score 54.2;  DB 3;  Length 1557;
  Best Local Similarity   55.6%;  Pred. No. 0.0021;
  Matches  104;  Conservative    0;  Mismatches   83;  Indels    0;  Gaps    0;

Qy    609 cttggccattgagatttcgcatcctgcatgaaattgcccttggtgtaaattacctgcaca 668
```

```
           | |||||  |    |  ||| | ||| | ||| | |  |||||  |   ||||||||||
Db     338 cctggccgctcctttgccgcctgctgaaagaagtggtgcttgggatgttttacctgcacg 397

Qy     669 atatgactcctcctttacttcatcatgacttgaagactcagaatatcttattggacaatg 728
           |   ||  ||     | || || |  ||| | ||| |    ||  || |  |||||   |
Db     398 accagaacccggtgctcctgcaccgggacctcaagccatccaacgtcctgctggacccag 457

Qy     729 aatttcatgttaagattgcagattttggtttatcaaagtggcgcatgatgtccctctcac 788
           |  | || || ||| | |||||||||||  | || |  |  |        || |  |||
Db     458 agctgcacgtcaagctggcagattttggcctgtccacatttcagggaggctcacagtcag 517

Qy     789 agtcacg 795
            | || |
Db     518 ggacagg 524

RESULT  12
US-09-329-418-1
; Sequence 1, Application US/09329418
; Patent No. 6096539
; GENERAL INFORMATION:
;  APPLICANT: ZENECA Limited
;  TITLE OF INVENTION: PROTEIN ACTIVATOR OF APOPTOSIS
;  FILE REFERENCE: PHM.70536
;  CURRENT APPLICATION NUMBER: US/09/329,418
;  CURRENT FILING DATE: 1999-06-11
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 1873
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-09-329-418-1


  Query Match             2.6%;  Score 54.2;  DB 3;  Length 1873;
  Best Local Similarity   55.6%;  Pred. No. 0.0022;
  Matches  104;  Conservative    0;  Mismatches   83;  Indels    0;  Gaps    0;

Qy     609 cttggccattgagatttcgcatcctgcatgaaattgcccttggtgtaaattacctgcaca 668
           | |||||  |    |  ||| | ||| | ||| | |  |||||  |   ||||||||||
Db     505 cctggccgctcctttgccgcctgctgaaagaagtggtgcttgggatgttttacctgcacg 564

Qy     669 atatgactcctcctttacttcatcatgacttgaagactcagaatatcttattggacaatg 728
           |   ||  ||     | || || |  ||| | ||| |    ||  || |  |||||   |
Db     565 accagaacccggtgctcctgcaccgggacctcaagccatccaacgtcctgctggacccag 624

Qy     729 aatttcatgttaagattgcagattttggtttatcaaagtggcgcatgatgtccctctcac 788
           |  | || || ||| | |||||||||||  | || |  |  |        || |  |||
Db     625 agctgcacgtcaagctggcagattttggcctgtccacatttcagggaggctcacagtcag 684

Qy     789 agtcacg 795
            | || |
Db     685 ggacagg 691

RESULT  13
US-08-545-196B-10
; Sequence 10, Application US/08545196B
; Patent No. 6080577
;  GENERAL INFORMATION:
;    APPLICANT:  MELKI, JUDITH
;    APPLICANT:  MUNNICH, ARNOLD
;    TITLE OF INVENTION:  SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
;    TITLE OF INVENTION:  FOR SPINAL MUSCULAR ATROPHY
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH AND BIRCH, LLP
;      STREET:  PO BOX 747
;      CITY:  FALLS CHURCH
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22040-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/545,196B
;      FILING DATE:  19-OCT-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FARACI, C. J.
;      REGISTRATION NUMBER:  32,350
;      REFERENCE/DOCKET NUMBER:  2121-110P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 205-8000
;      TELEFAX:  (703) 205-8050
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1582 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-545-196B-10


  Query Match             2.6%;  Score 53.6;  DB 3;  Length 1582;
  Best Local Similarity   49.2%;  Pred. No. 0.0029;
  Matches  122;  Conservative    6;  Mismatches  120;  Indels    0;  Gaps    0;

Qy    1850 tctttaaatttacttcaaaataaaagcatgtaagtgactgtttttcaagaagaaatgtgt 1909
           |||| |||| |     |   |  ||  |  |   ||   |   |    | . ||||||
Db    1335 TCTTAAAATGTTTCAAATGGTTTAACAAAATGTATGTGAGGCGTATGTGGCAAAATGTTA 1394

Qy    1910 ttcataaaaggatatttatatctctgttgctttgactttttttatataaaatccgtgagt 1969
              |    |     |  | ||  |||||  ||  ||| ||||| | ||   ||  |  ||
Db    1395 CAGAATCTAACTGGTGGACATGGCTGTTCATTGTACTGTTTTTTTCTATCTTCTATATGT 1454

Qy    1970 attaaagcttwawwraargktctttsrktaaatattagtctccctccatgacactgcagt 2029
            | ||||  |: :: ||:  | ||| ::|   | |||          |  | |    |
Db    1455 TTAAAAGTATATAATAAAAATATTTAATTTTTTTTTAAAAAAAAAAAAAAAAAAAAAAAA 1514

Qy    2030 attttttttaattaatacaagtaaaaagttgaatttgaaaaaaaaaaaaaaaaaaaaaaa 2089
           |        ||  || | ||  |||||    ||    |||||||||||||||||||||||
Db    1515 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1574

Qy    2090 aaaaaaaa 2097
           ||||||||
Db    1575 AAAAAAAA 1582


RESULT  14
US-08-545-196B-12
; Sequence 12, Application US/08545196B
; Patent No. 6080577
;  GENERAL INFORMATION:
;    APPLICANT:  MELKI, JUDITH
;    APPLICANT:  MUNNICH, ARNOLD
;    TITLE OF INVENTION:  SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
;    TITLE OF INVENTION:  FOR SPINAL MUSCULAR ATROPHY
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH AND BIRCH, LLP
;      STREET:  PO BOX 747
;      CITY:  FALLS CHURCH
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22040-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/545,196B
;      FILING DATE:  19-OCT-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FARACI, C. J.
;      REGISTRATION NUMBER:  32,350
;      REFERENCE/DOCKET NUMBER:  2121-110P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 205-8000
;      TELEFAX:  (703) 205-8050
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1582 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-545-196B-12.


  Query Match             2.6%;  Score 53.6;  DB 3;  Length 1582;
  Best Local Similarity   49.2%;  Pred. No. 0.0029;
  Matches  122;  Conservative    6;  Mismatches  120;  Indels    0;  Gaps    0;

Qy   1850 tctttaaatttacttcaaaataaaagcatgtaagtgactgtttttcaagaagaaatgtgt 1909
          |||| |||| |     |   |  ||  |  |   ||   |   |    |   ||||||
Db   1335 TCTTAAAATGTTTCAAATGGTTTAACAAAATGTATGTGAGGCGTATGTGGCAAAATGTTA 1394

Qy   1910 ttcataaaaggatatttatatctctgttgctttgactttttttatataaaatccgtgagt 1969
             |    |     |  | ||  |||||  ||  ||| ||||| | ||   ||  |  ||
Db   1395 CAGAATCTAACTGGTGGACATGGCTGTTCATTGTACTGTTTTTTTCTATCTTCTATATGT 1454

Qy   1970 attaaagcttwawwraargktctttsrktaaatattagtctccctccatgacactgcagt 2029
           | ||||  |: :: ||:  | ||| ::|   | |||          |  | |    |
Db   1455 TTAAAAGTATATAATAAAAATATTTAATTTTTTTTTAAAAAAAAAAAAAAAAAAAAAAAA 1514

Qy   2030 attttttttaattaatacaagtaaaaagttgaatttgaaaaaaaaaaaaaaaaaaaaaaa 2089
          |        ||  || | ||  |||||    ||    |||||||||||||||||||||||
Db   1515 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1574

Qy   2090 aaaaaaaa 2097
          ||||||||
Db   1575 AAAAAAAA 1582

RESULT  15
US-07-867-106-2/c
; Sequence 2, Application US/07867106
; Patent No. 5389526
;  GENERAL INFORMATION:
;    APPLICANT:  Slade, Martin B
;    APPLICANT:  Chang, Andy C M
;    APPLICANT:  Williams, Keith L
;    TITLE OF INVENTION:  Improved Plasmid Vectors for Cellular
;    TITLE OF INVENTION:  Slime Moulds of the Genus Dictyostelium
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
;      STREET:  One Liberty Place 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/867,106
;      FILING DATE:  19920625
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  AU PJ 7187
;      APPLICATION NUMBER:  PCT/AU90/00530
;      FILING DATE:  02-NOV-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Feeney, Joanne Longo
;      REGISTRATION NUMBER:  35,134
;      REFERENCE/DOCKET NUMBER:  RICE-0002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5852 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  2378..5038
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  2378..5038
US-07-867-106-2


  Query Match             2.5%;  Score 53.2;  DB 1;  Length 5852;
  Best Local Similarity   49.6%;  Pred. No. 0.006;
  Matches  121;  Conservative    5;  Mismatches  118;  Indels    0;  Gaps    0;

Qy   1855 aaatttacttcaaaataaaagcatgtaagtgactgtttttcaagaagaaatgtgtttcat 1914
          ||| | ||   |  |  ||   |  |   | |   || || |   | |  | |  |.  |
Db   5837 AAAGTAACAATAGTACTAATAAAGATCTTTTAAATTTATTAATATACATTTTTAATGGTT 5778

Qy   1915 aaaaggatatttatatctctgttgctttgactttttttatataaaatccgtgagtattaa 1974
          |      |  ||  ||   |  |  ||||  ||||||||||||   |   ||    ||
Db   5777 ATTTAATTTATTTAATTATTTGTTATTTGTATTTTTTTATATATGTTATTTGTTGTTTGT 5718

Qy   1975 agcttwawwraargktctttsrktaaatattagtctccctccatgacactgcagtatttt 2034
             ||:| : | : :||| |  :| | |  ||   |   |   | |   |  | ||  |
Db   5717 TTTTTTACTTATATTTCTATTTTTTATTTATAAAATTAATAATTAAATTTTAAATAAATA 5658

Qy   2035 ttttaattaatacaagtaaaaagttgaatttgaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2094
              ||  || | ||  ||||  |  | |   |||||||||||||||||||||||||| |
Db   5657 AAAAAAAAAAAAAAAAAAAAATTTAAATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAATA 5598

Qy   2095 aaaa 2098
          ||||
Db   5597 AAAA 5594


Search completed: June 13, 2001, 17:33:58
Job time: 6861 sec
```

```
                    GenCore version 4.5
           Copyright (c) 1993 - 2000  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         June 13, 2001, 11:10:10 ; Search time 1923.15 Seconds
                                          (without alignments)
                                          9530.371 Million cell updates/sec

Title:          US-09-445-223-2
Perfect score:  2098
Sequence:       1 ggccattatggatggatggg..........aaaaaaaaaaaaaaaaaaa 2098

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       9623517 seqs, 4368049070 residues

Total number of hits satisfying chosen parameters:      19247034

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       EST:*
               1:  gb_est1:*
               2:  gb_est2:*
               3:  gb_est3:*
               4:  gb_est4:*
               5:  gb_est5:*
               6:  gb_est6:*
               7:  gb_est7:*
               8:  gb_est8:*
               9:  gb_est9:*
               10:  gb_est10:*
               11:  gb_est11:*
               12:  gb_est12:*
               13:  gb_est13:*
               14:  gb_est14:*
               15:  gb_est15:*
               16:  gb_est16:*
               17:  gb_est17:*
               18:  gb_est18:*
               19:  gb_est19:*
               20:  gb_est20:*
               21:  gb_est21:*
               22:  gb_est22:*
               23:  gb_est23:*
               24:  gb_est24:*
               25:  gb_est33:*
               26:  gb_est34:*
               27:  gb_est35:*
               28:  gb_est36:*
               29:  gb_est37:*
               30:  gb_est38:*
               31:  gb_est39:*
               32:  gb_est40:*
               33:  em_estba:*
               34:  em_estfun:*
               35:  em_esthum1:*
               36:  em_esthum2:*
               37:  em_esthum3:*
               38:  em_esthum4:*
               39:  em_esthum5:*
               40:  em_esthum6:*
               41:  em_esthum7:*
               42:  em_esthum8:*
               43:  em_esthum9:*
               44:  em_esthum10:*
               45:  em_esthum11:*
               46:  em_esthum12:*
               47:  em_esthum13:*
               48:  em_esthum14:*
               49:  em_esthum15:*
               50:  em_esthum16:*
               51:  em_esthum17:*
               52:  em_esthum18:*
               53:  em_esthum19:*
               54:  em_esthum20:*
               55:  em_esthum21:*
               56:  em_esthum22:*
               57:  em_esthum23:*
               58:  em_esthum24:*
               59:  em_esthum25:*
               60:  em_esthum26:*
               61:  em_esthum27:*
               62:  em_esthum28:*
               63:  em_estin1:*
               64:  em_estin2:*
               65:  em_estin3:*
               66:  em_estin4:*
               67:  em_estin5:*
               68:  em_estom1:*
               69:  em_estom2:*
               70:  em_estov1:*
               71:  em_estov2:*
               72:  em_estpl1:*
               73:  em_estpl2:*
               74:  em_estpl3:*
               75:  em_estpl4:*
               76:  em_estpl5:*
               77:  em_estpl6:*
               78:  em_estpl7:*
               79:  em_estpl8:*
               80:  em_estpl9:*
               81:  em_estpl10:*
               82:  em_estro1:*
               83:  em_estro2:*
               84:  em_estro3:*
               85:  em_estro4:*
               86:  em_estro5:*
               87:  em_estro6:*
               88:  em_estro7:*
               89:  em_estro8:*
               90:  em_estro9:*
               91:  em_estro10:*
               92:  em_estro11:*
               93:  em_estro12:*
               94:  em_estro13:*
               95:  em_estro14:*
               96:  em_estro15:*
               97:  em_estro16:*
               98:  em_estro17:*
               99:  em_estro18:*
               100:  em_estro19:*
               101:  em_estro20:*
               102:  gb_est25:*
               103:  gb_est26:*
               104:  gb_est27:*
               105:  gb_est28:*
               106:  gb_est29:*
               107:  gb_est30:*
               108:  gb_est31:*
               109:  gb_est32:*
               110:  gb_est41:*
               111:  gb_est42:*
               112:  gb_est43:*
               113:  gb_est44:*
               114:  gb_est45:*
               115:  gb_est46:*
               116:  gb_est47:*
```

```
117:  gb_est48:*
118:  gb_est49:*
119:  gb_est50:*
120:  gb_est51:*
121:  gb_est52:*
122:  gb_est53:*
123:  gb_est54:*
124:  gb_est55:*
125:  gb_est56:*
126:  gb_est57:*
127:  gb_est58:*
128:  gb_est59:*
129:  gb_est60:*
130:  gb_est61:*
131:  gb_est62:*
132:  gb_est63:*
133:  gb_est64:*
134:  gb_est65:*
135:  gb_est66:*
136:  gb_est75:*
137:  gb_est76:*
138:  gb_est77:*
139:  gb_est78:*
140:  gb_est79:*
141:  gb_est80:*
142:  gb_est81:*
143:  gb_est82:*
144:  gb_est83:*
145:  gb_est84:*
146:  gb_est85:*
147:  gb_est86:*
148:  gb_est87:*
149:  gb_est88:*
150:  gb_est89:*
151:  gb_est90:*
152:  gb_est99:*
153:  gb_est100:*
154:  gb_est101:*
155:  gb_est102:*
156:  gb_est103:*
157:  gb_est104:*
158:  gb_est105:*
159:  gb_est106:*
160:  gb_est67:*
161:  gb_est68:*
162:  gb_est69:*
163:  gb_est70:*
164:  gb_est71:*
165:  gb_est72:*
166:  gb_est73:*
167:  gb_est74:*
168:  gb_est91:*
169:  gb_est92:*
170:  gb_est93:*
171:  gb_est94:*
172:  gb_est95:*
173:  gb_est96:*
174:  gb_est97:*
175:  gb_est98:*
176:  em_gss_fun:*
177:  em_gss_hum1:*
178:  em_gss_hum2:*
179:  em_gss_hum3:*
180:  em_gss_hum4:*
181:  em_gss_hum5:*
182:  em_gss_hum6:*
183:  em_gss_hum7:*
184:  em_gss_hum8:*
185:  em_gss_hum9:*
186:  em_gss_inv1:*
187:  em_gss_inv2:*
188:  em_gss_inv3:*
189:  em_gss_other:*
```

```
190:  em_gss_pln1:*
191:  em_gss_pln2:*
192:  em_gss_pro:*
193:  em_gss_rod1:*
194:  em_gss_rod2:*
195:  em_gss_rod3:*
196:  em_gss_rod4:*
197:  em_gss_rod5:*
198:  em_gss_vrt1:*
199:  em_gss_vrt2:*
200:  em_gss_vrt3:*
201:  gb_gss1:*
202:  gb_gss2:*
203:  gb_gss3:*
204:  gb_gss4:*
205:  gb_gss5:*
206:  gb_gss6:*
207:  gb_gss7:*
208:  gb_gss8:*
209:  gb_gss9:*
210:  gb_gss10:*
211:  gb_gss11:*
212:  gb_gss12:*
213:  gb_gss13:*
214:  gb_gss14:*
215:  gb_gss15:*
216:  gb_gss16:*
217:  gb_gss17:*
218:  gb_gss18:*
219:  gb_gss19:*
220:  gb_gss20:*
221:  gb_gss21:*
222:  gb_gss22:*
223:  gb_gss23:*
224:  gb_gss24:*
225:  gb_gss25:*
226:  gb_gss26:*
227:  gb_gss27:*
228:  gb_gss28:*
229:  gb_gss29:*
230:  gb_gss30:*
231:  gb_gss31:*
232:  gb_gss32:*
233:  gb_gss33:*
234:  gb_gss34:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                     Description
---------------------------------------------------------------------------
    1   745.8    35.5    811  174  BG170405              BG170405 602322736
    2     686    32.7    870  140  BE877822              BE877822 601486392
    3   655.8    31.3    828  140  BE875947              BE875947 601486423
c   4   621.8    29.6    636  136  BE551615              BE551615 7a42g06.x
c   5   609.4    29.0    641  111  AW150819              AW150819 xg39f08.x
c   6     592    28.2    592  24  AI745575              AI745575 wc34f12.x
c   7   590.6    28.2    606  110  AW085560              AW085560 wy67c04.x
c   8   588.8    28.1    647  122  AW960501              AW960501 EST372572
    9     587    28.0    654  136  BE536247              BE536247 601062632
c  10     561    26.7    615  102  AI801150              AI801150 to85h04.x
   11   558.6    26.6    645  3  AA160647              AA160647 zq49c11.r
   12     554    26.4    839  144  BF125028              BF125028 601762610
   13   550.4    26.2   1112  144  BF125423              BF125423 601763549
c  14   546.8    26.1    616  18  AI307810              AI307810 tb28d07.x
   15     542    25.8    762  103  AI904799              AI904799 IL-BT067-
c  16   532.6    25.4    568  3  AA161113              AA161113 zo58d05.s
   17   513.4    24.5    917  165  BE274455              BE274455 601120495
c  18     506    24.1    518  117  AW593657              AW593657 xl94d07.x
```

```
c  19   478.2   22.8    483  19  AI343247                        AI343247 tb94a04.x
   20   466.8   22.2    484  5  AA315575                        AA315575 EST187344
c  21     464   22.1    476  13  AA913804                        AA913804 ol35d11.s
c  22   452.2   21.6    487  114  AW393452                        AW393452 RC5-ST030
c  23   450.6   21.5    496  114  AW393431                        AW393431 RC5-ST030
c  24     449   21.4    455  19  AI343850                        AI343850 qp07d01.x
c  25   441.4   21.0    445  120  AW820715                        AW820715 RC5-ST030
c  26   439.8   21.0    445  120  AW820730                        AW820730 RC5-ST030
   27   438.2   20.9    940  169  BF796292                        BF796292 602258601
   28   430.4   20.5    446  117  AW610171                        AW610171 RC5-ST030
c  29   428.4   20.4    448  21  AI499817                        AI499817 tm92h03.x
c  30   425.6   20.3    439  19  AI380994                        AI380994 tg20c09.x
c  31     423   20.2    423  9  AA574167                        AA574167 nl78b01.s
c  32   422.8   20.2    446  161  BE002713                        BE002713 QV4-BN009
c  33     422   20.1    443  16  AI123747                        AI123747 oo21c04.x
c  34     418   19.9    432  11  AA723533                        AA723533 zg73g08.s
   35   416.6   19.9    472  168  BF706215                        BF706215 280591 MA
   36   410.2   19.6    503  10  AA655189                        AA655189 vv13a12.r
   37   402.4   19.2    426  114  AW385642                        AW385642 PMO-LT003
c  38     394   18.8    400  12  AA826052                        AA826052 od60b09.s
c  39   388.6   18.5    423  143  BF062840                        BF062840 7h70g03.x
   40   382.8   18.2    432  3  AA158827                        AA158827 zo58d05.r
c  41   382.8   18.2    462  114  AW361643                        AW361643 MR2-CT025
c  42   381.4   18.2    401  11  AA740822                        AA740822 ny98e05.s
c  43   371.2   17.7    493  175  C05866                        C05866 C05866 Huma
c  44   366.2   17.5    510  114  AW389863                        AW389863 RC2-ST017
   45     364   17.3    447  114  AW361603                        AW361603 MR2-CT025
```

ALIGNMENTS

```
RESULT   1
BG170405
LOCUS       BG170405       811 bp    mRNA            EST       06-FEB-2001
DEFINITION  602322736F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4426016 5',
            mRNA sequence.
ACCESSION   BG170405
VERSION     BG170405.1  GI:12677108
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 811)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10172  row: g  column: 09
            High quality sequence stop: 721.
FEATURES             Location/Qualifiers
     source          1. .811
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4426016"
                     /clone_lib="NIH_MGC_89"
                     /tissue_type="hypernephroma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
```

```
BASE COUNT      272 a    177 c    156 g    206 t
ORIGIN

  Query Match             35.5%;  Score 745.8;  DB 174;  Length 811;
  Best Local Similarity   96.9%;  Pred. No. 2e-153;
  Matches  782;  Conservative    0;  Mismatches   22;  Indels    3;  Gaps    2;

Qy   1032 atatacctcaccgagcacgtatgatctctctaatagaaagtggatgggcacaaaatccag 1091
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATATACCTCACCGAGCACGTATGATCTCTCTAATAGAAAGTGGATGGGCACAAAATCCAG 60

Qy   1092 atgaaagaccatctttcttaaaatgtttaatagaacttgaaccagttttgagaacatttg 1151
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATGAAAGACCATCTTTCTTAAAATGTTTAATAGAACTTGAACCAGTTTTGAGAACATTTG 120

Qy   1152 aagagataactttttcttgaagctgttattcagctaaagaaaacaaagttacagagtgttt 1211
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AAGAGATAACTTTTCTTGAAGCTGTTATTCAGCTAAAGAAAACAAAGTTACAGAGTGTTT 180

Qy   1212 caagtgccattcacctatgtgacaagaagaaaatggaattatctctgaacatacctgtaa 1271
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CAAGTGCCATTCACCTATGTGACAAGAAGAAAATGGAATTATCTCTGAACATACCTGTAA 240

Qy   1272 atcatggtccacaagaggaatcatgtggatcctctcagctccatgaaaatagtggttctc 1331
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ATCATGGTCCACAAGAGGAATCATGTGGATCCTCTCAGCTCCATGAAAATAGTGGTTCTC 300

Qy   1332 ctgaaacttcaaggtccctgccagctcctcaagacaatgatttttatctagaaaagctc 1391
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CTGAAACTTCAAGGTCCCTGCCAGCTCCTCAAGACAATGATTTTTATCTAGAAAAGCTC 360

Qy   1392 aagactgttattttatgaagctgcatcactgtcctggaaatcacagttgggatagcacca 1451
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AAGACTGTTATTTTATGAAGCTGCATCACTGTCCTGGAAATCACAGTTGGGATAGCACCA 420

Qy   1452 tttctggatctcaaagggctgcattctgtgatcacaagaccactccatgctcttcagcaa 1511
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TTTCTGGATCTCAAAGGGCTGCATTCTGTGATCACAAGACCACTCCATGCTCTTCAGCAA 480

Qy   1512 taataaatccactctcaactgcaggaaactcagaacgtctgcagcctggtatagcccagc 1571
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TAATAAATCCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAGCCTGGTATAGCCCAGC 540

Qy   1572 agtggatccagagcaaaagggaagacattgtgaaccaaatgacagaagcctgccttaacc 1631
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AGTGGATCCAGAGCAAAAGGGAAGACATTGTGAACCAAATGACAGAAGCCTGCCTTAACC 600

Qy   1632 agtcgctagatgcccttctgtccagggacttgatcatgaaagaggactatgaacttgtta 1691
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AGTCGCTAGATGCCCTTCTGTCCAGGGACTTGATCATGAAAGAGGACTATGAACTTGTTA 660

Qy   1692 gtaccaagcctacaaggacctcaaaagtcaga-caattactagacactactgacatccaa 1750
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db    661 GTACCAAGCCTACAAGGACCTCAAAAGTCAGACCAATTACTAGACACTACTGACATCCAA 720

Qy   1751 ggagaagaatttgccaaagttatagtacaaaaattgaaagataacaaacaaatgggtctt 1810
          |||||||| ||  | ||  |    |||||||||||||    || |||   ||||||||||
Db    721 GGAGAAGATTTGCCAAAGTTATACGTACAAAAATTGACGATTACCAA--CAATGGGTCTT 778

Qy   1811 cagccttacccggaaatacttgtggtt 1837
           ||||||||| | |||||||   ||||
Db    779 AAGCCTTACCGGAAAATACTGTGGGTT 805


RESULT   2
BE877822
LOCUS       BE877822       870 bp    mRNA            EST       20-OCT-2000
DEFINITION  601486392F1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:3888815 5',
            mRNA sequence.
ACCESSION   BE877822
VERSION     BE877822.1  GI:10326598
KEYWORDS    EST.
```

```
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 870)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9669  row: g  column: 24
            High quality sequence stop: 739.
FEATURES             Location/Qualifiers
     source          1..870
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3888815"
                     /clone_lib="NIH_MGC_69"
                     /tissue_type="large cell carcinoma, undifferentiated"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.1 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      288 a    200 c    170 g    212 t
ORIGIN


  Query Match             32.7%;  Score 686;  DB 140;  Length 870;
  Best Local Similarity   97.2%;  Pred. No. 2.7e-140;
  Matches  762;  Conservative    0;  Mismatches   15;  Indels    7;  Gaps    6;

Qy   1153 agagataacttttcttgaagctgttattcagctaaagaaaacaaagttacagagtgtttc 1212
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AGAGATAACTTTTCTTGAAGCTGTTATTCAGCTAAAGAAAACAAAGTTACAGAGTGTTTC 60

Qy   1213 aagtgccattcacctatgtgacaagaagaaaatggaattatctctgaacatacctgtaaa 1272
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AAGTGCCATTCACCTATGTGACAAGAAGAAAATGGAATTATCTCTGAACATACCTGTAAA 120

Qy   1273 tcatggtccacaagaggaatcatgtggatcctctcagctccatgaaaatagtggttctcc 1332
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TCATGGTCCACAAGAGGAATCATGTGGATCCTCTCAGCTCCATGAAAATAGTGGTTCTCC 180

Qy   1333 tgaaacttcaaggtccctgccagctcctcaagacaatgatttttatctagaaaagctca 1392
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TGAAACTTCAAGGTCCCTGCCAGCTCCTCAAGACAATGATTTTTATCTAGAAAAGCTCA 240

Qy   1393 agactgttattttatgaagctgcatcactgtcctggaaatcacagttgggatagcaccat 1452
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AGACTGTTATTTTATGAAGCTGCATCACTGTCCTGGAAATCACAGTTGGGATAGCACCAT 300

Qy   1453 ttctggatctcaaagggctgcattctgtgatcacaagaccactccatgctcttcagcaat 1512
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TTCTGGATCTCAAAGGGCTGCATTCTGTGATCACAAGACCACTCCATGCTCTTCAGCAAT 360

Qy   1513 aataaatccactctcaactgcaggaaactcagaacgtctgcagcctggtatagcccagca 1572
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AATAAATCCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAGCCTGGTATAGCCCAGCA 420

Qy   1573 gtggatccagagcaaaagggaagacattgtgaaccaaatgacagaagcctgccttaacca 1632
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGGATCCAGAGCAAAAGGGAAGACATTGTGAACCAAATGACAGAAGCCTGCCTTAACCA 480

Qy   1633 gtcgctagatgcccttctgtccagggacttgatcatgaaagaggactatgaacttgttag 1692
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    481 GTCGCTAGATGCCCTTCTGTCCA-GGACTTGATCATGAAAGAGGACTATGAACTTGTTAG 539

Qy   1693 taccaagcctacaaggacctcaaaagtcagacaattactagacactactgacatccaa-g 1751
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    540 TACCAAGCCTACAAGGACCTCAAAAGTCAGACAATTACTAGACACTACTGACATCCAAGG 599

Qy   1752 gagaagaatttgccaaagttatagtacaaaaattgaaagataacaaacaaatgggtctt- 1810
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 GAGAAGAATTTGCCAAAGTTATAGTACAAAAATTGAAAGATAACAAACAAATGGGTCTTT 659

Qy   1811 -cagccttacccggaaatacttgtggtttctagatcaccatctttaaa-tttacttcaaa 1868
             ||||||||||||||||||||||||||||||||||||||||||||| ||||||||| |
Db    660 CAGGCCTTACCCGGAAATACTTGTGGTTTCTAGATCACCATCTTTAAATTTACTTCAGA 719

Qy   1869 ataaaagcatgtaagtgactgtttttcaagaagaaatgtgtttcataaaaggatatttat 1928
          ||||||||   |||||||||||  ||| || ||| ||||| ||| | |  ||||||||||
Db    720 ATAAAAGCTGTTAAGTGACTGTGCTTC-AGCAGACATGTG-TTCCTTACCGGATATTTAT 777

Qy   1929 atct 1932
           |||
Db    778 CTCT 781


RESULT   3
BE875947
LOCUS       BE875947        828 bp    mRNA            EST       20-OCT-2000
DEFINITION  601486423F1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:3888965 5',
            mRNA sequence.
ACCESSION   BE875947
VERSION     BE875947.1  GI:10324723
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 828)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9669  row: n  column: 06
            High quality sequence stop: 795.
FEATURES             Location/Qualifiers
     source          1..828
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3888965"
                     /clone_lib="NIH_MGC_69"
                     /tissue_type="large cell carcinoma, undifferentiated"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.1 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      287 a    163 c    165 g    213 t
ORIGIN


  Query Match             31.3%;  Score 655.8;  DB 140;  Length 828;
  Best Local Similarity   98.4%;  Pred. No. 1.2e-133;
  Matches  694;  Conservative    0;  Mismatches    7;  Indels    4;  Gaps    3;

Qy   1267 tgtaaatcatggtccacaagaggaatcatgtggatcctctcagctccatgaaaatagtgg 1326
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4 TGTAAATCATGGTCCACAAGAGGAATCATGTGGATCCTCTCAGCTCCATGAAAATAGTGG 63

Qy   1327 ttctcctgaaacttcaaggtccctgccagctcctcaagacaatgattttttatctagaaa 1386
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64 TTCTCCTGAAACTTCAAGGTCCCTGCCAGCTCCTCAAGACAATGATTTTTTATCTAGAAA 123

Qy   1387 agctcaagactgttattttatgaagctgcatcactgtcctggaaatcacagttgggatag 1446
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124 AGCTCAAGACTGTTATTTTATGAAGCTGCATCACTGTCCTGGAAATCACAGTTGGGATAG 183

Qy   1447 caccatttctggatctcaaagggctgcattctgtgatcacaagaccactccatgctcttc 1506
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184 CACCATTTCTGGATCTCAAAGGGCTGCATTCTGTGATCACAAGACCACTCCATGCTCTTC 243

Qy   1507 agcaataataaatccactctcaactgcaggaaactcagaacgtctgcagcctggtatagc 1566
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 AGCAATAATAAATCCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAGCCTGGTATAGC 303

Qy   1567 ccagcagtggatccagagcaaaagggaagacattgtgaaccaaatgacagaagcctgcct 1626
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 CCAGCAGTGGATCCAGAGCAAAAGGGAAGACATTGTGAACCAAATGACAGAAGCCTGCCT 363

Qy   1627 taaccagtcgctagatgcccttctgtccagggacttgatcatgaaagaggactatgaact 1686
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    364 TAACCAGTCGCTAGATGCCCTTCTGTCCA-GGACTTGATCATGAAAGAGGACTATGAACT 422

Qy   1687 tgttagtaccaagcctacaaggacctcaaaagtcagacaattactagacactactgacat 1746
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    423 TGTTAGTACCAAGCCTACAAGGACCTCAAAAGTCAGACAATTACTAGACACTACTGACAT 482

Qy   1747 ccaaggagaagaatttgccaaagttatagtacaaaaattgaaagataacaaacaaatggg 1806
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    483 CCAAGGAGAAGAATTTGCCAAAGTTATAGTACAAAAATTGAAAGATAACAAACAAATGGG 542

Qy   1807 tcttcagccttacccggaaatactt-gtggtttctagatcaccatctttaaatttacttc 1865
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    543 TCTTCAGCCTTACCCGGAAATACTTGGTGGTTTCTAGATCACCATCTTTAAATTTACTTC 602

Qy   1866 aaaataaaagcatgtaagtgactgtttttcaagaagaaatgtgtttcataaaaggatatt 1925
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    603 AAAATAAAAGCATGTAAGTGACTGTTTTTCAAGAAGAAATGTGTTTCATAAAAGGATATT 662

Qy   1926 tatatctctgttgctttgactttttttatataaaatccgtgagta 1970
          |||||||||| ||  |||||||||||||||||||  | | |  ||
Db    663 TATATCTCTGGTG--TTGACTTTTTTTATATAAATCCGGAGTATA 705


RESULT   4
BE551615/c
LOCUS       BE551615       636 bp    mRNA            EST       10-AUG-2000
DEFINITION  7a42g06.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:3221434 3'
            similar to TR:O43353 O43353 SERINE/THREONINE KINASE RICK. ;, mRNA
            sequence.
ACCESSION   BE551615
VERSION     BE551615.1  GI:9793227
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 636)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
```

```
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 455.
FEATURES             Location/Qualifiers
     source          1..636
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3221434"
                     /clone_lib="NCI_CGAP_GC6"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Plasmid DNA from the normalized library
                     NCI_CGAP_GC4 was prepared, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from a pool of 5,000 clones made
                     from the same library (cloneIDs 1257096-1258631,
                     1469064-1470983, and 1475592-1476743). Subtraction by
                     Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      175 a    118 c    135 g    208 t
ORIGIN


  Query Match             29.6%;  Score 621.8;  DB 136;  Length 636;
  Best Local Similarity   99.5%;  Pred. No. 3.2e-126;
  Matches  634;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy   1337 acttcaaggtccctgccagctcctcaagacaatgattttttatctagaaaagctcaagac 1396
          |||||| ||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db    636 ACTTCAGGGTCCCTGCCAGCTCCTCAAGACAATG-TTTTTTATCTAGAAAAGCTCAAGAC 578

Qy   1397 tgttattttatgaagctgcatcactgtcctggaaatcacagttgggatagcaccatttct 1456
          |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db    577 TGTTATTTTATGAAGCTGCATCCCTGTCCTGGAAATCACAGTTGGGATAGCACCATTTCT 518

Qy   1457 ggatctcaaagggctgcattctgtgatcacaagaccactccatgctcttcagcaataata 1516
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    517 GGATCTCAAAGGGCTGCATTCTGTGATCACAAGACCACTCCATGCTCTTCAGCAATAATA 458

Qy   1517 aatccactctcaactgcaggaaactcagaacgtctgcagcctggtatagcccagcagtgg 1576
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    457 AATCCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAGCCTGGTATAGCCCAGCAGTGG 398

Qy   1577 atccagagcaaaagggaagacattgtgaaccaaatgacagaagcctgccttaaccagtcg 1636
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    397 ATCCAGAGCAAAAGGGAAGACATTGTGAACCAAATGACAGAAGCCTGCCTTAACCAGTCG 338

Qy   1637 ctagatgcccttctgtccagggacttgatcatgaaagaggactatgaacttgttagtacc 1696
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    337 CTAGATGCCCTTCTGTCCAGGGACTTGATCATGAAAGAGGACTATGAACTTGTTAGTACC 278

Qy   1697 aagcctacaaggacctcaaaagtcagacaattactagacactactgacatccaaggagaa 1756
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    277 AAGCCTACAAGGACCTCAAAAGTCAGACAATTACTAGACACTACTGACATCCAAGGAGAA 218

Qy   1757 gaatttgccaaagttatagtacaaaaattgaaagataacaaacaaatgggtcttcagcct 1816
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    217 GAATTTGCCAAAGTTATAGTACAAAAATTGAAAGATAACAAACAAATGGGTCTTCAGCCT 158

Qy   1817 tacccggaaatacttgtggtttctagatcaccatctttaaatttacttcaaaataaaagc 1876
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    157 TACCCGGAAATACTTGTGGTTTCTAGATCACCATCTTTAAATTTACTTCAAAATAAAAGC 98

Qy   1877 atgtaagtgactgtttttcaagaagaaatgtgtttcataaaaggatatttatatctctgt 1936
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     97 ATGTAAGTGACTGTTTTTCAAGAAGAAATGTGTTTCATAAAAGGATATTTATATCTCTGT 38

Qy   1937 tgctttgactttttttatataaaatccgtgagtatta 1973
```

```
          |||||||||||||||||||||||||||||||||||||
Db     37 TGCTTTGACTTTTTTTATATAAAATCCGTGAGTATTA 1


RESULT   5
AW150819/c
LOCUS       AW150819      641 bp    mRNA            EST       03-NOV-1999
DEFINITION  xg39f08.x1 NCI_CGAP_Ut1 Homo sapiens cDNA clone IMAGE:2629959 3'
            similar to TR:O43353 O43353 SERINE/THREONINE KINASE RICK. ;, mRNA
            sequence.
ACCESSION   AW150819
VERSION     AW150819.1  GI:6198717
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 641)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco
            High quality sequence stop: 402.
FEATURES             Location/Qualifiers
     source          1..641
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2629959"
                     /clone_lib="NCI_CGAP_Ut1"
                     /tissue_type="well-differentiated endometrial
                     adenocarcinoma, 7 pooled tumors"
                     /lab_host="DH10B"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb. Life Technologies catalog #:
                     11538-014"
BASE COUNT      185 a    116 c    123 g    215 t      2 others
ORIGIN


  Query Match             29.0%;  Score 609.4;  DB 111;  Length 641;
  Best Local Similarity   95.9%;  Pred. No. 1.7e-123;
  Matches  613;  Conservative    9;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   1427 ggaaatcacagttgggatagcaccatttctggatctcaaagggctgcattctgtgatcac 1486
          |||||||| ||| || |||||||||||| ||||| ||| |||||||| |||||| |||||
Db    641 GGAAATCAAAGTGGGAATAGCACCATTTNTGGATGTCACAGGGCTGCTTTCTGTAATCAC 582

Qy   1487 aagaccactccatgctcttcagcaataataaatccactctcaactgcaggaaactcagaa 1546
          ||| |||  ||| |||||||||||||||||||||||||||||||||||| ||||||||||
Db    581 AAGCCCATCCCACGCTCTTCAGCAATAATAAATCCACTCTCAACTGCAGNAAACTCAGAA 522

Qy   1547 cgtctgcagcctggtatagcccagcagtggatccagagcaaaagggaagacattgtgaac 1606
          |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db    521 CGTCTGCAGCTTGGTATAGCCCAGCAGTGGATCCAGAGCAAAAGGGAAGACATTGTGAAC 462

Qy   1607 caaatgacagaagcctgccttaaccagtcgctagatgcccttctgtccagggacttgatc 1666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    461 CAAATGACAGAAGCCTGCCTTAACCAGTCGCTAGATGCCCTTCTGTCCAGGGACTTGATC 402

Qy   1667 atgaaagaggactatgaacttgttagtaccaagcctacaaggacctcaaaagtcagacaa 1726
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    401 ATGAAAGAGGACTATGAACTTGTTAGTACCAAGCCTACAAGGACCTCAAAAGTCAGACAA 342

Qy   1727 ttactagacactactgacatccaaggagaagaatttgccaaagttatagtacaaaaattg 1786
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    341 TTACTAGACACTACTGACATCCAAGGAGAAGAATTTGCCAAAGTTATAGTACAAAAATTG 282

Qy   1787 aaagataacaaacaaatgggtcttcagccttacccggaaatacttgtggtttctagatca 1846
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    281 AAAGATAACAAACAAATGGGTCTTCAGCCTTACCCGGAAATACTTGTGGTTTCTAGATCA 222

Qy   1847 ccatctttaaatttacttcaaaataaaagcatgtaagtgactgtttttcaagaagaaatg 1906
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221 CCATCTTTAAATTTACTTCAAAATAAAAGCATGTAAGTGACTGTTTTTCAAGAAGAAATG 162

Qy   1907 tgtttcataaaaggatatttatatctctgttgctttgactttttttatataaaatccgtg 1966
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db.   161 TGTTTCATAAAAGGATATTTATATCTCTGTTGCTTTGACTTTTTTTATATAAAATCCGTG 102

Qy   1967 agtattaaagcttwawwraargktctttsrktaaatattagtctccctccatgacactgc 2026
          |||||||||||||:|:::||:|:|||||:::|||||||||||||||||||||||||||||
Db    101 AGTATTAAAGCTTTATTGAAGGTTCTTTGGGTAAATATTAGTCTCCCTCCATGACACTGC 42

Qy   2027 agtattttttttaattaatacaagtaaaaagttgaattt 2065
          |||||||||||||||||||||||||||||||||  | ||
Db     41 AGTATTTTTTTTAATTAATACAAGTAAAAAGTTTGAATT 3


RESULT   6
AI745575/c
LOCUS       AI745575      592 bp    mRNA            EST       17-DEC-1999
DEFINITION  wc34f12.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2317103 3'
            similar to TR:O43353 O43353 SERINE/THREONINE KINASE RICK. ;, mRNA
            sequence.
ACCESSION   AI745575
VERSION     AI745575.1  GI:5113863
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 592)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 673   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 459.
FEATURES             Location/Qualifiers
     source          1..592
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2317103"
                     /clone_lib="NCI_CGAP_Pr28"
                     /sex="male"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Plasmid DNA from the
                     normalized library NCI_CGAP_Pr22 was prepared, and ss
                     circles were made in vitro. Following HAP purification,
```

```
                     this DNA was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from a pool
                     of 5,000 clones made from the same library (cloneIDs
                     985608-986759, 1101192-1101959, and 1217928-1220615).
                     Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      153 a    111 c    128 g    200 t
ORIGIN

  Query Match             28.2%;  Score 592;  DB 24;  Length 592;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-119;
  Matches  592;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1337 acttcaaggtccctgccagctcctcaagacaatgatttttatctagaaaagctcaagac 1396
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    592 ACTTCAAGGTCCCTGCCAGCTCCTCAAGACAATGATTTTTATCTAGAAAAGCTCAAGAC 533

Qy   1397 tgttattttatgaagctgcatcactgtcctggaaatcacagttgggatagcaccatttct 1456
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    532 TGTTATTTTATGAAGCTGCATCACTGTCCTGGAAATCACAGTTGGGATAGCACCATTTCT 473

Qy   1457 ggatctcaaagggctgcattctgtgatcacaagaccactccatgctcttcagcaataata 1516
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    472 GGATCTCAAAGGGCTGCATTCTGTGATCACAAGACCACTCCATGCTCTTCAGCAATAATA 413

Qy   1517 aatccactctcaactgcaggaaactcagaacgtctgcagcctggtatagcccagcagtgg 1576
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    412 AATCCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAGCCTGGTATAGCCCAGCAGTGG 353

Qy   1577 atccagagcaaaagggaagacattgtgaaccaaatgacagaagcctgccttaaccagtcg 1636
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    352 ATCCAGAGCAAAAGGGAAGACATTGTGAACCAAATGACAGAAGCCTGCCTTAACCAGTCG 293

Qy   1637 ctagatgcccttctgtccagggacttgatcatgaaagaggactatgaacttgttagtacc 1696
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    292 CTAGATGCCCTTCTGTCCAGGGACTTGATCATGAAAGAGGACTATGAACTTGTTAGTACC 233

Qy   1697 aagcctacaaggacctcaaaagtcagacaattactagacactactgacatccaaggagaa 1756
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    232 AAGCCTACAAGGACCTCAAAAGTCAGACAATTACTAGACACTACTGACATCCAAGGAGAA 173

Qy   1757 gaatttgccaaagttatagtacaaaaattgaaagataacaaacaaatgggtcttcagcct 1816
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    172 GAATTTGCCAAAGTTATAGTACAAAAATTGAAAGATAACAAACAAATGGGTCTTCAGCCT 113

Qy   1817 tacccggaaatacttgtggtttctagatcaccatctttaaatttacttcaaaataaaagc 1876
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    112 TACCCGGAAATACTTGTGGTTTCTAGATCACCATCTTTAAATTTACTTCAAAATAAAAGC 53

Qy   1877 atgtaagtgactgtttttcaagaagaaatgtgtttcataaaaggatatttat 1928
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     52 ATGTAAGTGACTGTTTTTCAAGAAGAAATGTGTTTCATAAAAGGATATTTAT 1

RESULT    7
AW085560/c
LOCUS       AW085560       606 bp    mRNA            EST       09-MAR-2000
DEFINITION  wy67c04.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
            IMAGE:2553606 3' similar to TR:O43353 O43353 SERINE/THREONINE
            KINASE RICK. ;, mRNA sequence.
ACCESSION   AW085560
VERSION     AW085560.1  GI:6040712
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 606)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 885   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 470.
FEATURES             Location/Qualifiers
     source          1. .606
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2553606"
                     /clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from five normalized
                     libraries were mixed, and ss circles were made in vitro.
                     Following HAP purification, this DNA was used as tracer in
                     a subtractive hybridization reaction. The driver was
                     PCR-amplified cDNAs from pools of 5,000 clones made from
                     the same 5 libraries. The pools consisted of the following
                     libraries and cloneIDs: Soares NbHSF pool 1:
                     309384-310919, 323208-325895 Soares Nb2HP pool 1:
                     145032-147335, 147720-148103, 148872-149255, 15002 -
                     150407, 151176-152327 Soares Nb2HF8-9W pool 1:
                     758280-760583, 772104-774407 Soares NbHPA pool 1:
                     304776-306311, 320136-322823, 326280-326663 Soares NbHOT
                     pool 1:  723720-726407, 739080-740999 Subtraction by Bento
                     Soares and M. Fatima Bonaldo."
BASE COUNT      166 a    114 c    122 g    203 t      1 others
ORIGIN

  Query Match             28.2%;  Score 590.6;  DB 110;  Length 606;
  Best Local Similarity   98.7%;  Pred. No. 2.3e-119;
  Matches  598;  Conservative    6;  Mismatches    1;  Indels    1;  Gaps    1;

Qy   1385 aaagctcaagactgttattttatgaagctgcatcactgtcctggaaatcacagttgggat 1444
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    606 AAAGCTCAAGACTGTTATTNTATGAAGCTGCATCACTGTCCTGGAAATCACAGTTGGGAT 547

Qy   1445 agcaccatttctggatctcaaagggctgcattctgtgatcacaagaccactccatgctct 1504
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    546 AGCACCATTTCTGGATCTCAAAGGGCTGCATTCTGTGATCACAAGACCACTCCATGCTCT 487

Qy   1505 tcagcaataataaatccactctcaactgcaggaaactcagaacgtctgcagcctggtata 1564
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    486 TCAGCAATAATAAATCCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAGCCTGGTATA 427

Qy   1565 gcccagcagtggatccagagcaaaagggaagacattgtgaaccaaatgacagaagcctgc 1624
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    426 GCCCAGCAGTGGATCCAGAGCAAAAGGGAAGACATTGTGAACCAAATGACAGAAGCCTGC 367

Qy   1625 cttaaccagtcgctagatgcccttctgtccagggac-ttgatcatgaaagaggactatga 1683
          |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db    366 CTTAACCAGTCGCTAGATGCCCTTCTGTCCAGGGACTTTGATCATGAAAGAGGACTATGA 307

Qy   1684 acttgttagtaccaagcctacaaggacctcaaaagtcagacaattactagacactactga 1743
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    306 ACTTGTTAGTACCAAGCCTACAAGGACCTCAAAAGTCAGACAATTACTAGACACTACTGA 247

Qy   1744 catccaaggagaagaatttgccaaagttatagtacaaaaattgaaagataacaaacaaat 1803
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    246 CATCCAAGGAGAAGAATTTGCCAAAGTTATAGTACAAAAATTGAAAGATAACAAACAAAT 187

Qy   1804 gggtcttcagccttacccggaaatacttgtggtttctagatcaccatctttaaatttact 1863
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    186 GGGTCTTCAGCCTTACCCGGAAATACTTGTGGTTTCTAGATCACCATCTTTAAATTTACT 127

Qy   1864 tcaaaataaaagcatgtaagtgactgtttttcaagaagaaatgtgtttcataaaaggata 1923
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    126 TCAAAATAAAAGCATGTAAGTGACTGTTTTTCAAGAAGAAATGTGTTTCATAAAAGGATA 67
```

```
Qy   1924 tttatatctctgttgctttgactttttttatataaaatccgtgagtattaaagcttwaww 1983
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|::
Db     66 TTTATATCTCTGTTGCTTTGACTTTTTTTATATAAAATCCGTGAGTATTAAAGCTTTATT 7

Qy   1984 raargk 1989
          :||:|:
Db      6 GAAGGT 1


RESULT   8
AW960501/c
LOCUS       AW960501       647 bp    mRNA             EST       01-JUN-2000
DEFINITION  EST372572 MAGE resequences, MAGF Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW960501
VERSION     AW960501.1  GI:8150185
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 647)
  AUTHORS   Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt
            ,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
            Quackenbush,J.
  TITLE     Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
  JOURNAL   Unpublished (2000)
COMMENT     Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            Plate: 147
            Seq primer: Forward.
FEATURES             Location/Qualifiers
     source          1. .647
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="MAGE resequences, MAGF"
                     /note="Vector: pBluescriptSKm"
BASE COUNT      165 a    128 c    141 g    213 t
ORIGIN


  Query Match            28.1%;  Score 588.8;  DB 122;  Length 647;
  Best Local Similarity  98.7%;  Pred. No. 5.7e-119;
  Matches  604;  Conservative    0;  Mismatches    7;  Indels    1;  Gaps    1;

Qy   1317 aaaatagtggttctcctgaaacttcaa-ggtccctgccagctcctcaagacaatgatttt 1375
          ||||||||||||||| | ||| ||||| |||||||||||||||  |||| ||||||||||
Db    612 AAAATAGTGGTTCTCTTAAAATTTCAAGGGTCCCTGCCAGCTCTCCAAGCCAATGATTTT 553

Qy   1376 ttatctagaaaagctcaagactgttattttatgaagctgcatcactgtcctggaaatcac 1435
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    552 TTATCTAGAAAAGCTCAAGCCTGTTATTTTATGAAGCTGCATCACTGTCCTGGAAATCAC 493

Qy   1436 agttgggatagcaccatttctggatctcaaagggctgcattctgtgatcacaagaccact 1495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    492 AGTTGGGATAGCACCATTTCTGGATCTCAAAGGGCTGCATTCTGTGATCACAAGACCACT 433

Qy   1496 ccatgctcttcagcaataataaatccactctcaactgcaggaaactcagaacgtctgcag 1555
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    432 CCATGCTCTTCAGCAATAATAAATCCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAG 373

Qy   1556 cctggtatagcccagcagtggatccagagcaaaagggaagacattgtgaaccaaatgaca 1615
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    372 CCTGGTATAGCCCAGCAGTGGATCCAGAGCAAAAGGGAAGACATTGTGAACCAAATGACA 313

Qy   1616 gaagcctgccttaaccagtcgctagatgcccttctgtccagggacttgatcatgaaagag 1675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    312 GAAGCCTGCCTTAACCAGTCGCTAGATGCCCTTCTGTCCAGGGACTTGATCATGAAAGAG 253
```

```
Qy   1676 gactatgaacttgttagtaccaagcctacaaggacctcaaaagtcagacaattactagac 1735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    252 GACTATGAACTTGTTAGTACCAAGCCTACAAGGACCTCAAAAGTCAGACAATTACTAGAC 193

Qy   1736 actactgacatccaaggagaagaatttgccaaagttatagtacaaaaattgaaagataac 1795
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    192 ACTACTGACATCCAAGGAGAAGAATTTGCCAAAGTTATAGTACAAAAATTGAAAGATAAC 133

Qy   1796 aaacaaatgggtcttcagccttacccggaaatacttgtggtttctagatcaccatcttta 1855
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    132 AAACAAATGGGTCTTCAGCCTTACCCGGAAATACTTGTGGTTTCTAGATCACCATCTTTA 73

Qy   1856 aatttacttcaaaataaaagcatgtaagtgactgttttcaagaagaaatgtgtttcata 1915
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     72 AATTTACTTCAAAATAAAAGCATGTAAGTGACTGTTTTCAAGAAGAAATGTGTTTCATA 13

Qy   1916 aaaggatattta 1927
          ||||||||||||
Db     12 AAAGGATATTTA 1


RESULT   9
BE536247
LOCUS       BE536247       654 bp    mRNA             EST       09-AUG-2000
DEFINITION  601062632F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:3449311 5',
            mRNA sequence.
ACCESSION   BE536247
VERSION     BE536247.1  GI:9764892
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 654)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Incyte Genomics, Inc.
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM8425  row: o  column: 08
            High quality sequence stop: 650.
FEATURES             Location/Qualifiers
     source          1. .654
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3449311"
                     /clone_lib="NIH_MGC_10"
                     /cell_line="MGC36"
                     /lab_host="DH10B"
                     /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.5 kb. Library prepared by Life
                     Technologies."
BASE COUNT      214 a    149 c    131 g    158 t      2 others
ORIGIN


  Query Match            28.0%;  Score 587;  DB 136;  Length 654;
  Best Local Similarity  97.6%;  Pred. No. 1.4e-118;
  Matches  638;  Conservative    0;  Mismatches   11;  Indels    5;  Gaps    4;

Qy   1188 agaaaacaaagttacagagtgtttcaagtgccattcacctatgtgacaagaagaaaatgg 1247
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AGAAAACAAAGTTACAGAGTGTTTCAAGTGCCATTCACCTATGTGACAAGAAGAAAATGG 60
```

```
Qy   1248 aattatctctgaacatacctgtaaatcatggtccacaagaggaatcatgtggatcctctc 1307
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AATTATCTCTGAACATACCTGTAAATCATGGTCCACAAGAGGAATCATGTGGATCCTCTC 120

Qy   1308 a-gctccatgaaaatagtggttctcctgaaacttcaaggtccctgccagctcctcaagac 1366
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ANGCTCCATGAAAATAGTGGTTCTCCTGAAACTTCAAGGTCCCTGCCAGCTCCTCAAGAC 180

Qy   1367 aatgattttttatctagaaaagctcaagactgttattttatgaagctgcatcactgtcct 1426
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AATGATTTTTTATCTAGAAAAGCTCAAGACTGTTATTTTATGAAGCTGCATCACTGTCCT 240

Qy   1427 ggaaatcacagttgggatagcaccatttctggatctcaaagggctgcattctgtgatcac 1486
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGAAATCACAGTTGGGATAGCACCATTTCTGGATCTCAAAGGGCTGCATTCTGTGATCAC 300

Qy   1487 aagaccactccatgctcttcagcaataataaatccactctcaactgcaggaaactcagaa 1546
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AAGACCACTCCATGCTCTTCAGCAATAATAAATCCACTCTCAACTGCAGGAAACTCAGAA 360

Qy   1547 cgtctgcagcctggtatagcccagcagtggatccagagcaaaagggaagacattgtgaac 1606
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CGTCTGCAGCCTGGTATAGCCCAGCAGTGGATCCAGAGCAAAAGGGAAGACATTGTGAAC 420

Qy   1607 caaatgacagaagcctgccttaaccagtcgctagatgcccttctgtccagggacttgatc 1666
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    421 CAAATGACAGAAGCCTG-CTTAACCAGTCGCTAGATGCCCTTCTGTCCAGGGACTTGATC 479

Qy   1667 atgaaagaggactatgaacttgttagtaccaagcctacaaggacctcaaaagtcagacaa 1726
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 ATGAAAGAGGACTATGAACTTGTTAGTACCAAGCCTACAAGGACCTCAAAAGTCAGACAA 539

Qy   1727 ttactagacactactgacatccaagg-agaagaatttgccaaagttatagtacaaaaatt 1785
          || ||||||||||||||||||||||| ||||||||||| || ||||||||||||||||||
Db    540 TTGCTAGACACTACTGACATCCAAGGCAGAAGAATTTGGCANAGTTATAGTACAAAAATT 599

Qy   1786 --gaaagataacaaacaaatgggtcttcagccttacccggaaatacttgtggtt 1837
             ||    | ||||||||||||||||||||||||||||| ||||||||||| |
Db    600 TGCAAGATAACCAAACAAATGGGTCTTCAGCCTTACCCGGGAATACTTGTGGCT 653


RESULT  10
AI801150/c
LOCUS       AI801150      615 bp    mRNA            EST       14-DEC-1999
DEFINITION  to85h04.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2185111 3'
            similar to TR:O43353 O43353 SERINE/THREONINE KINASE RICK. ;, mRNA
            sequence.
ACCESSION   AI801150
VERSION     AI801150.1  GI:5366622
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 615)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 3507   Std Error: 0.00
            Seq primer: -40UP from Gibco
```

```
                High quality sequence stop: 421.
FEATURES             Location/Qualifiers
     source          1..615
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2185111"
                     /clone_lib="NCI_CGAP_Gas4"
                     /tissue_type="poorly differentiated adenocarcinoma with
                     signet ring cell features"
                     /lab_host="DH10B"
                     /note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.69 kb. Life Technologies catalog #:
                     11549-011"
BASE COUNT      179 a    110 c    120 g    206 t
ORIGIN


  Query Match            26.7%;  Score 561;  DB 102;  Length 615;
  Best Local Similarity  97.4%;  Pred. No. 7.2e-113;
  Matches  591;  Conservative    9;  Mismatches    4;  Indels    3;  Gaps    3;

Qy   1460 tctcaaagggctgcattctgtgatcacaagaccactccatgctcttcagcaataataaat 1519
          |||||||||||||||||||||||| ||| ||||||||||||||||||||||||||| |||
Db    615 TCTCAAAGGGCTGCATTCTGTGAT-ACAGGACCACTCCATGCTCTTCAGCAATAAT-AAT 558

Qy   1520 ccactctcaactgcaggaaactcagaacgtctgcag-cctggtatagcccagcagtggat 1578
          |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db    557 CCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAGCCCTGGTATAGCCCAGCAGTGGAT 498

Qy   1579 ccagagcaaaagggaagacattgtgaaccaaatgacagaagcctgccttaaccagtcgct 1638
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    497 CCAGAGCAAAAGGGAAGACATTGTGAACCAAATGACAGAAGCCTGCCTTAACCAGTCGCT 438

Qy   1639 agatgcccttctgtccagggacttgatcatgaaagaggactatgaacttgttagtaccaa 1698
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    437 AGATGCCCTTCTGTCCAGGGACTTGATCATGAAAGAGGACTATGAACTTGTTAGTACCAA 378

Qy   1699 gcctacaaggacctcaaaagtcagacaattactagacactactgacatccaaggagaaga 1758
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    377 GCCTACAAGGACCTCAAAAGTCAGACAATTACTAGACACTACTGACATCCAAGGAGAAGA 318

Qy   1759 atttgccaaagttatagtacaaaaattgaaagataacaaacaaatgggtcttcagcctta 1818
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    317 ATTTGCCAAAGTTATAGTACAAAAATTGAAAGATAACAAACAAATGGGTCTTCAGCCTTA 258

Qy   1819 cccggaaatacttgtggtttctagatcaccatctttaaatttacttcaaaataaaagcat 1878
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    257 CCCGGAAATACTTGTGGTTTCTAGATCACCATCTTTAAATTTACTTCAAAATAAAAGCAT 198

Qy   1879 gtaagtgactgtttttcaagaagaaatgtgtttcataaaaggatatttatatctctgttg 1938
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    197 GTAAGTGACTGTTTTTCAAGAAGAAATGTGTTTCATAAAAGGATATTTATATCTCTGTTG 138

Qy   1939 ctttgactttttttatataaaatccgtgagtattaaagcttwawwraargktctttsrkt 1998
          |||||||||||||||||||||||||||||||||||||||||:|:::||:|:|||||:::|
Db    137 CTTTGACTTTTTTTATATAAAATCCGTGAGTATTAAAGCTTTATTGAAGGTTCTTTGGGT 78

Qy   1999 aaatattagtctccctccatgacactgcagtatttttttaattaatacaagtaaaagt 2058
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     77 AAATATTAGTCTCCCTCCATGACACTGCAGTATTTTTTTAATTAATACAAGTAAAAGT 18

Qy   2059 tgaattt 2065
          |  | ||
Db     17 TTGAATT 11


RESULT  11
AA160647
LOCUS       AA160647      645 bp    mRNA            EST       16-DEC-1996
DEFINITION  zq49c11.r1 Stratagene hNT neuron (#937233) Homo sapiens cDNA clone
            IMAGE:633044 5', mRNA sequence.
ACCESSION   AA160647
```

```
VERSION     AA160647.1  GI:1736024
KEYWORDS    EST.
SOURCE      human..
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 645)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 373.
FEATURES             Location/Qualifiers
     source          1..645
                     /organism="Homo sapiens"
                     /db_xref="GDB:5183128"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:633044"
                     /clone_lib="Stratagene hNT neuron (#937233)"
                     /dev_stage="hNT neurons"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT.
                     Differentiated, post mitotic hNT neurons. Average insert
                     size: 1.5 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5'
                     GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTTTT 3'"
BASE COUNT      205 a    138 c    123 g    178 t      1 others
ORIGIN


  Query Match             26.6%;  Score 558.6;  DB 3;  Length 645;
  Best Local Similarity   99.1%;  Pred. No. 2.4e-112;
  Matches  572;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy     952 tgtcaccaatcctttgcagataatgtatagtgtgtcacaaggacatcgacctgttattaa 1011
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TGTCACCAATCCTTTGCAGATAATGTATAGTGTGTCACAAGGACATCGACCTGTTATTAA 60

Qy    1012 tgaagaaagtttgccatatgatatacctcaccgagcacgtatgatctctctaatagaaag 1071
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TGAAGAAAGTTTGCCATATGATATACCTCACCGAGCACGTATGATCTCTCTAATAGAAAG 120

Qy    1072 tggatgggcacaaaatccagatgaaagaccatctttcttaaaatgtttaatagaacttga 1131
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TGGATGGGCACAAAATCCAGATGAAAGACCATCTTTCTTAAAATGTTTAATAGAACTTGA 180

Qy    1132 accagttttgagaacatttgaagagataacttttcttgaagctgttattcagctaaagaa 1191
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ACCAGTTTTGAGAACATTTGAAGAGATAACTTTTCTTGAAGCTGTTATTCAGCTAAAGAA 240

Qy    1192 aacaaagttacagagtgtttcaagtgccattcacctatgtgacaagaagaaaatggaatt 1251
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AACAAAGTTACAGAGTGTTTCAAGTGCCATTCACCTATGTGACAAGAAGAAAATGGAATT 300

Qy    1252 atctctgaacatacctgtaaatcatggtccacaagaggaatcatgtggatcctctcagct 1311
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ATCTCTGAACATACCTGTAAATCATGGTCCACAAGAGGAATCATGTGGATCCTCTCAGCT 360
```

```
Qy    1312 ccatgaaaatagtggttctcctgaaacttcaaggtccctgccagctcctcaagacaatga 1371
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CCATGAAAATAGTGGTTCTCCTGAAACTTCAAGGTCCCTGCCAGCTCCTCAAGACAATGA 420

Qy    1372 tttttatctagaaaagctcaagactgttattttatgaagctgcatcactgtcctggaaa 1431
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TTTTTATCTAGAAAAGCTCAAGACTGTTATTTTATGAAGCTGCATCACTGTCCTGGAAA 480

Qy    1432 tcacagttgggatagcaccatttctggatctcaaagggctgcattctgtgatcacaagac 1491
           ||||||||||||||||||||||||||||||||||||||||||||||||||||    ||||
Db     481 TCACAGTTGGGATAGCACCATTTCTGGATCTCAAAGGGCTGCATTCTGTGATTCACAGAC 540

Qy    1492 cactccatgctcttcagcaataataaatccactctca 1528
           ||||||||||||||||| |||||||||||||||||||
Db     541 CACTCCATGCTCTTCAGC-ATAATAAATCCACTCTCA 576


RESULT  12
BF125028
LOCUS       BF125028      839 bp    mRNA            EST       24-OCT-2000
DEFINITION  601762610F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4025427 5',
            mRNA sequence.
ACCESSION   BF125028
VERSION     BF125028.1  GI:10964068
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 839)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM854  row: d  column: 04
            High quality sequence stop: 636.
FEATURES             Location/Qualifiers
     source          1..839
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4025427"
                     /clone_lib="NIH_MGC_20"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      213 a    195 c    246 g    185 t
ORIGIN


  Query Match             26.4%;  Score 554;  DB 144;  Length 839;
  Best Local Similarity   94.8%;  Pred. No. 2.6e-111;
  Matches  660;  Conservative    0;  Mismatches   25;  Indels   11;  Gaps    8;

Qy      77 agcagcggctggcgtgggccatccggggaatgggcgccctcgtgacctagtgttgcgggg 136
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 AGCAGCGGCTGGCGTGGGCCATCCGGGGAATGGGCGCCCTCGTGACCTAGTGTTGCGGGG 60

Qy     137 caaaaagggtcttgccggcctcgctcgtgcaggggcgtatctgggcgcctgagcgcggcg 196
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CAAAAAGGGTCTTGCCGGCCTCGCTCGTGCAGGGGCGTATCTGGGCGCCTGAGCGCGGCG 120

Qy    197 tgggagccttgggagccgccgcagcaggggcacacccggaaccggcctgagcgcccggg 256
          |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db    121 TGGGAGCCTTGGGAGCCGCCGCAGCAGGGGGCACACCCGGAACCGG-CTGAGCGCCCGGG 179

Qy    257 accatgaacggggaggccatctgcagcgccctgcccaccattccctaccacaaactcgcc 316
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 ACCATGAACGGGGAGGCCATCTGCAGCGCCCTGCCCACCATTCCCTACCACAAACTCGCC 239

Qy    317 gacctgcgctacctgagccgcggcgcctctggcactgtgtcgtccgcccgccacgcagac 376
          |||||||||||||||||||| ||||||||||||||||||||||||| |||||||||||||
Db    240 GACCTGCGCTACCTGAGCCG-GGCGCCTCTGGCACTGTGTCGTCCG-CCGCCACGCAGAC 297

Qy    377 tggcgcgtccaggtggccgtgaagcacctgcacatccacactccgctgctcgacagtga- 435
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    298 TGGCGCGTCCAGGTGGCCGTGAAGCACCTGCACATCCACACTCCGCTGCTCGACAGTGAC 357

Qy    436 aagaaaggatgttttaagagaagctgaaatttacacaaagctagatttagttacatttt 495
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||| |
Db    358 AAGAAAGGATGTCTTAAGAGAAGCTGAAATTTTACACAAAGCTAGATTTAGTTACATTCT 417

Qy    496 tccaattttgggaatttgcaatgagcctgaatttttgggaatagttactgaatacatgcc 555
          |||||||| || | ||   | ||||||||||||| |||||||||||||||||||||||||
Db    418 TCCAATTTGGGAATTTGCAACTGAGCCTGAATTTGTGGGAATAGTTACTGAATACATGCC 477

Qy    556 aaatggatcattaaatgaactcctacataggaaaactgaatatcctgatgttgcttggc- 614
          ||||||||||||| ||||||||||||||||||||||||||||||||||||| |||||||
Db    478 AAATGGATCATTACATGAACTCCTACATAGGAAAACTGAATATCCTGATGTGGCTTGGCG 537

Qy    615 cattgagatttcgcatc-ctgcatgaaattgcccttggtgtaaattacctgcacaatatg 673
          ||||||||||||||||| |||||||||||||||  |||||||||||||||| ||||||||
Db    538 CATTGAGATTTCGCATCGCTGCATGAAATTGCC--TGGTGTAAATTACCTGGACAATATG 595

Qy    674 actcctcctttacttcatcatgacttgaagactcagaatatcttattggacaatgaattt 733
          ||||||||||||||| ||      |||||||||||||||||||||||||||||||| |||
Db    596 ACTCCTCCTTTACTTTATTCATGATTGAAGACTCAGAATATCTTATTGGACAATGACTTT 655

Qy    734 catgttaagattgcagattttggtttatcaaagtgg 769
          |||||||||||    ||  |||| ||||||||||||
Db    656 CATGTTAAGAT---GGAGATTGGGTTATCAAAGTGG 688


RESULT  13
BF125423
LOCUS       BF125423      1112 bp    mRNA            EST       24-OCT-2000
DEFINITION  601763549F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4026156 5',
            mRNA sequence.
ACCESSION   BF125423
VERSION     BF125423.1  GI:10964463
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1112)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM856  row: b  column: 13
            High quality sequence stop: 566.
FEATURES             Location/Qualifiers
     source          1..1112
```

```
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4026156"
                     /clone_lib="NIH_MGC_20"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      305 a    312 c    294 g    200 t      1 others
ORIGIN


  Query Match             26.2%;  Score 550.4;  DB 144;  Length 1112;
  Best Local Similarity   89.1%;  Pred. No. 1.6e-110;
  Matches  675;  Conservative    0;  Mismatches   71;  Indels   12;  Gaps    7;

Qy     77 agcagcggctggcgtgggccatccggggaatgggcgccctcgtgacctagtgttgcgggg 136
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AGCAGCGGCTGGCGTGGGCCATCCGGGGAATGGGCGCCCTCGTGACCTAGTGTTGCGGGG 60

Qy    137 caaaaagggtcttgccggcctcgctcgtgcaggggcgtatctgggcgcctgagcgcggcg 196
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CAAAAAGGGTCTTGCCGGCCTCGCTCGTGCAGGGGCGTATCTGGGCGCCTGAGCGCGGCG 120

Qy    197 tgggagccttgggagccgccgcagcaggggcacacccggaaccggcctgagcgcccggg 256
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGGGAGCCTTGGGAGCCGCCGCAGCAGGGGGCACACCCGGAACCGGCCTGAGCGCCCGGG 180

Qy    257 accatgaacggggaggccatctgcagcgccctgcccaccattccctaccacaaactcgcc 316
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACCATGAACGGGGAGGCCATCTGCAGCGCCCTGCCCACCATTCCCTACCACAAACTCGCC 240

Qy    317 gacctgcgctacctgagccgcggcgcctctggcactgtgtcgtccgcccgccacgcagac 376
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GACCTGCGCTACCTGAGCCGCGGCGCCTCTGGCACTGTGTCGTCCGCCCGCCACGCAGAC 300

Qy    377 tggcgcgtccaggtggccgtgaagcacctgcacatccacactccgctgctcgacagtgaa 436
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGGCGCGTCCAGGTGGCCGTGAAGCACCTGCACATCCACACTCCGCTGCTCGACAGTGAA 360

Qy    437 agaaaggatgttttaagagaagctgaaatttacacaaagctagatttagttacatttt 496
          ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||| ||
Db    361 AGAAAGGATGTCTTAAGAGAAGCTGAAATTTTACACAAAGCTAGATTTAGTTACATTCTT 420

Qy    497 ccaattttgggaatttgcaatgagcctgaatttttgggaatagttactgaatacatgcca 556
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db    421 CCAATTTTGGGAATTTGCAATGAGCCTGAA-TTTTGGGAATAGTTACTGAATACATGCCA 479

Qy    557 aatggatcattaaatgaactcctacataggaaaactgaatatcctgatgttgcttggcca 616
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 AATGGATCATTAAATGAACTCCTACATAGGAAAACTGAATATCCTGATGTTGCTTGGC-- 537

Qy    617 ttgagatttcgcatcctgcatgaaattgcccttggtgtaaattacctgcacaatatgact 676
           |     ||||||||||| |||||||||||  |||||||| |||||||  ||||||||||
Db    538 ATTGAGATTCGCATCCTG-ATGAAATTGCC--TGGTGTAATTTACCTGACCAATATGACT 594

Qy    677 cctcctttacttcatcatgacttgaagactc--agaatatcttattggacaatgaatttc 734
          |   | |   |||||||||||||||||||||  | |    ||    || || || |  |
Db    595 C--TCCTGTATTCATCATGACTTGAAGACTCCGACATATCCTATGGGGCCACCGACTCGC 652

Qy    735 atgttaagattgcagattttggtttatcaaagtggcgcatgatgtccctctcacagtcac 794
          | || ||||| || | ||  | | |    | |||| |  ||||| ||     || | |
Db    653 AGGTAAAGATGGCGGCTTCGGCTATGCACACGTGGGGACTGATGGCC--TCTACGGCCCG 710

Qy    795 gaagtagcaaatctgcaccagaaggagggacaattatt 832
           | || || |||     |||||| | ||||  ||  ||
Db    711 CACGTCGCGAATTCTGCCCAGAATGGGGGAGCATAGTT 748
```

```
RESULT  14
AI307810/c
LOCUS       AI307810     616 bp    mRNA            EST       08-APR-1999
DEFINITION  tb28d07.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2055661 3',
            mRNA sequence.
ACCESSION   AI307810
VERSION     AI307810.1  GI:4002414
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 616)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 680   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 444.
FEATURES             Location/Qualifiers
     source          1. .616
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2055661"
                     /clone_lib="NCI_CGAP_Kid12"
                     /tissue_type="2 pooled tumors (clear cell type)"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid5 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1323912-1325831, 1471368-1472903 and
                     1492104-1493255). Subtraction by Bento Soares and M.
                     Fatima Bonaldo. "
BASE COUNT      178 a    109 c    112 g    217 t
ORIGIN


  Query Match             26.1%;  Score 546.8;  DB 18;  Length 616;
  Best Local Similarity   96.4%;  Pred. No. 9.4e-110;
  Matches  559;  Conservative    9;  Mismatches   11;  Indels    1;  Gaps    1;

Qy   1502 tcttcagcaataataaatccactctcaactgcaggaaactcagaacgtctgcagcctggt 1561
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    579 TATTCAGCAATAATAAATCCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAGCCTGGT 520

Qy   1562 atagcccagcagtggatccagagcaaaagggaagacattgtgaaccaaatgacagaagcc 1621
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    519 ATAGCCCAGCAGTGGATCCAGAGCAAAAGGGAAGACATTGTGAACCAAATGACAGAAGCC 460

Qy   1622 tgccttaaccagtcgctagatgcccttctgtccagggacttgatcatgaaagaggactat 1681
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    459 TGCCTTAACCAGTCGCTAGATGCCCTTCTGTCCAGGGACTTGATCATGAAAGAGGACTAT 400

Qy   1682 gaacttgttagtaccaagcctacaaggacctcaaaagtcagacaattactagacactact 1741
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    399 GAACTTGTTAGTACCAAGCCTACAAGGACCTCAAAAGTCAGACAATTACTAGACACTACT 340

Qy   1742 gacatccaaggagaagaatttgccaaagttatagtacaaaaattgaaagataacaaacaa 1801
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    339 GACATCCAAGGAGAAGAATTTGCCAAAGTTA-AGTACAAAAATTGAAAGATAACAAACAA 281

Qy   1802 atgggtcttcagccttacccggaaatacttgtggtttctagatcaccatctttaaattta 1861
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    280 ATGGGTCTTCAGCCTTACCCGGAAATACTTGTGGTTTCTAGATCACCATCTTTAAATTTA 221

Qy   1862 cttcaaaataaaagcatgtaagtgactgtttttcaagaagaaatgtgtttcataaaagga 1921
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    220 CTTCAAAATAAAAGCATGTAAGTGACTGTTTTTCAAGAAGAAATGTGTTTCATAAAAGGA 161

Qy   1922 tatttatatctctgttgctttgactttttttatataaaatccgtgagtattaaagcttwa 1981
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db    160 TATTTATATCTCTGTTGCTTTGACTTTTTTTATATAAAATCCGTGAGTATTAAAGCTTTA 101

Qy   1982 wwraargktctttsrktaaatattagtctccctccatgacactgcagtatttttttttaat 2041
          :::||:|:||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db    100 TTGAAGGTTCTTTGGGTAAATATTAGTCTCCCTCCATGACACTGCAGTATTTTTTTTTAAT 41

Qy   2042 taatacaagtaaaaagttgaatttgaaaaaaaaaaaaaaa 2081
          ||||||||||||||||||  | ||     | | |||||||
Db     40 TAATACAAGTAAAAAGTTTGAATTTTGCTACATAAAAAAA 1


RESULT  15
AI904799
LOCUS       AI904799     762 bp    mRNA            EST       30-MAR-2000
DEFINITION  IL-BT067-190199-023 BT067 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AI904799
VERSION     AI904799.1  GI:6495186
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 762)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/seq/gethtml.pl?t1=IL&t2=IL-BT067-023.html
            &t3=190199&t4=1)
            Seq primer: puc 18 forward.
FEATURES             Location/Qualifiers
     source          1. .762
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="BT067"
                     /sex="female"
                     /dev_stage="Adult"
                     /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No. 196
                     ,716 -  Ludwig Institute for Cancer Research) profiles
```

```
                    into  the pUC 18 vector. Reverse-transcription of tissue
                    mRNA and cDNA amplification were performed under low
                    stringency conditions."
BASE COUNT      232 a    157 c    137 g    208 t     28 others
ORIGIN


  Query Match             25.8%;  Score 542;  DB 103;  Length 762;
  Best Local Similarity   90.3%;  Pred. No. 1.1e-108;
  Matches  621;  Conservative    0;  Mismatches   62;  Indels    5;  Gaps    5;

Qy    931 cagaaaacagccttttgaagatgtcaccaatcctttgcagataatgtatagtgtgtcaca 990
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db      1 CAGAAAACAGCCTTTTGAAGATGTCACCAATCCTTTGCAGATAATGTATAGTGNGTCACA 60

Qy    991 aggacatcgacctgttattaatgaagaaagtttgccatatgatatacctcaccgagcacg 1050
           || ||||| |||| ||||||||||||||||||||||||||||| |||||||||||||||
Db     61 NGGNCATCGNCCTGNTATTAATGAAGAAAGTTTGCCATATGATACACCTCACCGAGCACG 120

Qy   1051 tatgatctctctaatagaaagtggatgggcacaaaatccagatgaaagaccatctttctt 1110
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TATGATCTCTCTAATAGAAAGTGGATGGGCACAAAATCCAGATGAAAGACCATCTTTCTT 180

Qy   1111 aaaatgtttaatagaacttgaaccagttttgagaacatttgaagagataactttcttga 1170
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAAATGTTTAATAGAACTTGAACCAGTTTTGAGAACATTTGAAGAGATAACTTTTCTTGA 240

Qy   1171 agctgttattcagctaaagaaaacaaagttacagagtgtttcaagtgccattcacctatg 1230
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AGCTGTTATTCAGCTAAAGAAAACAAAGTTACAGAGTGTTTCAAGTGCCATTCACCTATG 300

Qy   1231 tgacaagaagaaaatggaattatctctgaacatacctgtaaatcatggtccacaagagga 1290
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGACAAGAAGAAAATGGAATTATCTCTGAACATACCTGTAAATCATGGTCCACAAGAGGA 360

Qy   1291 atcatgtggatcctctcagctccatgaaaatagtggttctcctgaaacttcaaggtccct 1350
          |||||||||||||||||||||||||||||||||||| ||| |||||||||||||||||||
Db    361 ATCATGTGGATCCTCTCAGCTCCATGAAAATAGTGGGTCTTCTGAAACTTCAAGGTCCCT 420

Qy   1351 gccagctcctcaagacaatga-ttttttatctagaaaagctcaagactgttattttatga 1409
          ||||||||||||||||||||| |||||||| || ||||||||||||||||| |||| |||
Db    421 GCCAGCTCCTCAAGACAATGATTTTTTTATNTANAAAAGCTCAAGACTGTTNTTTT-TGA 479

Qy   1410 agctgcatcactgtcctggaaatcacagttgggatagcaccatttctggatctcaaaggg 1469
          | ||||||| ||||||||| |||| ||||||||||| |||||||| ||||| |  || ||
Db    480 ACCTGCATC-CTGTCCTGGNAATCGCAGTTGGGATA-CACCATTTTTGGATTTNNAAAGG 537

Qy   1470 ctgcattctgtgatcacaagaccactccatgctcttcagcaataataaatccactctcaa 1529
           |  ||| || |||||||| ||||||||||||| || || |||||||||||| || ||||
Db    538 GTTNATTTTGGGATCACAANACCACTCCATGCTNTTAAG-AATAATAAATCCCCTTTCAA 596

Qy   1530 ctgcaggaaactcagaacgtctgcagcctggtatagcccagcagtggatccagagcaaaa 1589
          |||  || |||||| ||    | | ||| ||  |  |||  | | || |||  ||| |||
Db    597 CTGGCGGNAACTCAAAAACGTTTCNGCCCGGGNTTACCCNCCCGNGGGTCCCAAGCNAAA 656

Qy   1590 gggaagacattgtgaaccaaatgacaga 1617
          ||||| |  ||  |||||||   | | |
Db    657 GGGAAAAANTTTGGAACCAANGAAAAAA 684


Search completed: June 13, 2001, 16:33:08
Job time: 19378 sec
```

```
                    GenCore version 4.5
          Copyright (c) 1993 - 2000  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         June 13, 2001, 15:58:26 ; Search time 169.91 Seconds
                                          (without alignments)
                                          7208.364 Million cell updates/sec

Title:          US-09-445-223-2
Perfect score:  2098
Sequence:       1 ggccattatggatggatggg..........aaaaaaaaaaaaaaaaaaaa 2098

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       678276 seqs, 291890651 residues

Total number of hits satisfying chosen parameters:      1356552

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      N_Geneseq_0401:*
               1:  /cgn1_8/gcgdata/geneseq/geneseqn/NA1980.DAT:*
               2:  /cgn1_8/gcgdata/geneseq/geneseqn/NA1981.DAT:*
               3:  /cgn1_8/gcgdata/geneseq/geneseqn/NA1982.DAT:*
               4:  /cgn1_8/gcgdata/geneseq/geneseqn/NA1983.DAT:*
               5:  /cgn1_8/gcgdata/geneseq/geneseqn/NA1984.DAT:*
               6:  /cgn1_8/gcgdata/geneseq/geneseqn/NA1985.DAT:*
               7:  /cgn1_8/gcgdata/geneseq/geneseqn/NA1986.DAT:*
               8:  /cgn1_8/gcgdata/geneseq/geneseqn/NA1987.DAT:*
               9:  /cgn1_8/gcgdata/geneseq/geneseqn/NA1988.DAT:*
               10:  /cgn1_8/gcgdata/geneseq/geneseqn/NA1989.DAT:*
               11:  /cgn1_8/gcgdata/geneseq/geneseqn/NA1990.DAT:*
               12:  /cgn1_8/gcgdata/geneseq/geneseqn/NA1991.DAT:*
               13:  /cgn1_8/gcgdata/geneseq/geneseqn/NA1992.DAT:*
               14:  /cgn1_8/gcgdata/geneseq/geneseqn/NA1993.DAT:*
               15:  /cgn1_8/gcgdata/geneseq/geneseqn/NA1994.DAT:*
               16:  /cgn1_8/gcgdata/geneseq/geneseqn/NA1995.DAT:*
               17:  /cgn1_8/gcgdata/geneseq/geneseqn/NA1996.DAT:*
               18:  /cgn1_8/gcgdata/geneseq/geneseqn/NA1997.DAT:*
               19:  /cgn1_8/gcgdata/geneseq/geneseqn/NA1998.DAT:*
               20:  /cgn1_8/gcgdata/geneseq/geneseqn/NA1999.DAT:*
               21:  /cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:*
               22:  /cgn1_8/gcgdata/geneseq/geneseqn/NA2001.DAT:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                              SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                  Description
-------------------------------------------------------------------------
   1   2094.4    99.8   2098  20  X02558               Human B1 cDNA.  Ho
   2   2031.6    96.8   2709  21  C77779               Human cancer assoc
   3   2006.8    95.7   2502  21  Z48762               Human RICK coding
   4   1999.6    95.3   2024  21  Z46143               cDNA sequence enco
   5     1864    88.8   1931  20  Z09246               Human CARD-3 cDNA.
   6    454.6    21.7    463  21  F21841               Human breast and o
   7      210    10.0    210  19  V33547               Clone 156108 cDNA
   8    102.4     4.9   1774  21  Z61830               cDNA encoding muri
   9    102.4     4.9   1774  22  C99763               Skin cell cDNA, SE
  10    102.4     4.9   2370  21  Z61161               DNA encoding a dea
  11    102.4     4.9   3516  21  Z61784               cDNA encoding muri
  12    102.4     4.9   3516  22  C99717               Skin cell cDNA, SE
  13      102     4.9    108  16  T19776               Human gene signatu
  14       93     4.4   2294  21  Z58584               Mouse protein kina
  15     81.4     3.9   3033  16  Q98816               Arabidopsis CTR1 c
  16     81.4     3.9   3033  18  T67126               Constitutive tripl
  17     79.8     3.8   2890  15  Q57911               Arabidopsis thalia
  18     78.2     3.7   1888  21  Z61671               cDNA encoding muri
  19     78.2     3.7   1888  22  C99604               Skin cell cDNA, SE
  20     66.4     3.2    667  21  A79568               Pinus radiata cell
  21     62.4     3.0    360  21  A79593               Pinus radiata cell
  22     61.8     2.9   1554  17  T31307               Tomato RRK gene cl
  23     61.8     2.9   3293  20  X23532               Tomato Xa21 clone
  24       60     2.9   2342  21  C38223               Arabidopsis thalia
  25       60     2.9   2505  21  A47607               Human CARK (Cardia
  26       60     2.9   2508  22  F44702               Novel protein kina
  27       60     2.9   3025  21  A47606               Human CARK (Cardia
  28     59.2     2.8   1322  21  F15889               Human prostate can
  29     59.2     2.8   3807  20  X37743               Human PRO326 DNA f
  30     59.2     2.8   4053  20  X52266               Protein PRO326 cDN
  31     59.2     2.8   4053  21  Z52208               Human PRO326 prote
  32     58.8     2.8    522  21  A79414               Eucalyptus grandis
  33     58.4     2.8   9789  17  T41852               cDNA encoding Plas
  34       58     2.8   1558  12  Q10896               Encodes Xenopus Bo
  35     56.4     2.7   2741  21  C47379               Arabidopsis thalia
  36     54.8     2.6    854  21  A79462               Eucalyptus grandis
  37     54.8     2.6   1433  21  C37768               Arabidopsis thalia
  38     54.2     2.6   1342  21  A26424               Human secreted pro
  39     54.2     2.6   1557  21  A47702               Kinase of death (K
  40     54.2     2.6   1873  21  A47701               Kinase of death (K
  41     54.2     2.6   2140  21  A75675               DNA encoding a hum
  42       54     2.6   2740  21  C36697               Arabidopsis thalia
  43     53.8     2.6   1864  19  V59630               Human secreted pro
  44     53.8     2.6   6002  20  X60227               Poly(A) binding pr
  45     53.6     2.6   1582  17  T28259               Survival motor neu

                              ALIGNMENTS

RESULT   1
X02558
ID   X02558 standard; cDNA; 2098 BP.
XX
AC   X02558;
XX
DT   07-MAY-1999  (first entry)
XX
DE   Human B1 cDNA.
XX
KW   B1 protein; intracellular mediator; modulator; inflammation; cell death;
KW   cell survival pathway; intracellular signalling; AIDS; cancer; human; ss.
XX
OS   Homo sapiens.
XX
PN   WO9855507-A2.
XX
PD   10-DEC-1998.
XX
PF   01-JUN-1998;   98WO-IL00255.
XX
PR   11-SEP-1997;   97IL-0121746.
PR   05-JUN-1997;   97IL-0121011.
PR   30-JUN-1997;   97IL-0121199.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Boldin M,  Malinin N,  Wallach D;
XX
DR   WPI; 1999-070258/06.
DR   P-PSDB; W92795.
XX
PT   New B1 protein regulates cell death and cell survival pathways -
PT   derivatives, DNA and antibodies, also regulate intracellular
```

```
PT   inflammation; for treating AIDS, cancer
XX
PS   Claim 4; Fig 3B; 90pp; English.
XX
CC   This invention describes the isolation of a novel human B1 protein which
CC   can interact with, intracellular mediators or modulators of inflammation,
CC   cell death and/or cell survival pathways, directly or indirectly. Cells
CC   can be modulated or mediated in inflammation, cell death or cell survival
CC   pathways or another intracellular signalling activity using B1.
CC   Conditions such as AIDS and cancer can be treated using B1. Antibodies,
CC   oligonucleotides and ribozymes can also be used to regulate the above
CC   pathways.
XX
SQ   Sequence 2098 BP; 649 A; 452 C; 449 G; 539 T; 9 other;

  Query Match            99.8%;  Score 2094.4;  DB 20;  Length 2098;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2098;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ggccattatggatggatgggcggcgctacggcgttggcaccagtctctagaaaagaagtc 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ggccattatggatggatgggcggcgctacggcgttggcaccagtctctagaaaagaagtc 60

Qy         61 agctctggttcggagaagcagcggctggcgtgggccatccggggaatgggcgccctcgtg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 agctctggttcggagaagcagcggctggcgtgggccatccggggaatgggcgccctcgtg 120

Qy        121 acctagtgttgcggggcaaaaagggtcttgccggcctcgctcgtgcaggggcgtatctgg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 acctagtgttgcggggcaaaaagggtcttgccggcctcgctcgtgcaggggcgtatctgg 180

Qy        181 gcgcctgagcgcggcgtgggagccttgggagccgccgcagcaggggcacacccggaacc 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 gcgcctgagcgcggcgtgggagccttgggagccgccgcagcaggggcacacccggaacc 240

Qy        241 ggcctgagcgcccgggaccatgaacggggaggccatctgcagcgccctgcccaccattcc 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ggcctgagcgcccgggaccatgaacggggaggccatctgcagcgccctgcccaccattcc 300

Qy        301 ctaccacaaactcgccgacctgcgctacctgagccgcggcgcctctggcactgtgtcgtc 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ctaccacaaactcgccgacctgcgctacctgagccgcggcgcctctggcactgtgtcgtc 360

Qy        361 cgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctgcacatccacactcc 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 cgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctgcacatccacactcc 420

Qy        421 gctgctcgacagtgaaagaaaggatgttttaagagaagctgaaattttacacaaagctag 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 gctgctcgacagtgaaagaaaggatgttttaagagaagctgaaattttacacaaagctag 480

Qy        481 atttagttacattttccaattttgggaatttgcaatgagcctgaatttttgggaatagt 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 atttagttacattttccaattttgggaatttgcaatgagcctgaatttttgggaatagt 540

Qy        541 tactgaatacatgccaaatggatcattaaatgaactcctacataggaaaactgaatatcc 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 tactgaatacatgccaaatggatcattaaatgaactcctacataggaaaactgaatatcc 600

Qy        601 tgatgttgcttggccattgagatttcgcatcctgcatgaaattgcccttggtgtaaatta 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 tgatgttgcttggccattgagatttcgcatcctgcatgaaattgcccttggtgtaaatta 660

Qy        661 cctgcacaatatgactcctcctttacttcatcatgacttgaagactcagaatatcttatt 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 cctgcacaatatgactcctcctttacttcatcatgacttgaagactcagaatatcttatt 720

Qy        721 ggacaatgaatttcatgttaagattgcagattttggtttatcaaagtggcgcatgatgtc 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ggacaatgaatttcatgttaagattgcagattttggtttatcaaagtggcgcatgatgtc 780

Qy        781 cctctcacagtcacgaagtagcaaatctgcaccagaaggagggacaattatttatatgcc 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 cctctcacagtcacgaagtagcaaatctgcaccagaaggagggacaattatttatatgcc 840

Qy        841 acctgaaaactatgaacctggacaaaaatcaagggccagtatcaagcacgatatatatag 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 acctgaaaactatgaacctggacaaaaatcaagggccagtatcaagcacgatatatatag 900

Qy        901 ctatgcagttatcacatgggaagtgttatccagaaaacagccttttgaagatgtcaccaa 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ctatgcagttatcacatgggaagtgttatccagaaaacagccttttgaagatgtcaccaa 960

Qy        961 tcctttgcagataatgtatagtgtgtcacaaggacatcgacctgttattaatgaagaaag 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 tcctttgcagataatgtatagtgtgtcacaaggacatcgacctgttattaatgaagaaag 1020

Qy       1021 tttgccatatgatatacctcaccgagcacgtatgatctctctaatagaaagtggatgggc 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 tttgccatatgatatacctcaccgagcacgtatgatctctctaatagaaagtggatgggc 1080

Qy       1081 acaaaatccagatgaaagaccatctttcttaaaatgtttaatagaacttgaaccagtttt 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 acaaaatccagatgaaagaccatctttcttaaaatgtttaatagaacttgaaccagtttt 1140

Qy       1141 gagaacatttgaagagataacttttcttgaagctgttattcagctaaagaaaacaaagtt 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 gagaacatttgaagagataacttttcttgaagctgttattcagctaaagaaaacaaagtt 1200

Qy       1201 acagagtgtttcaagtgccattcacctatgtgacaagaagaaaatggaattatctctgaa 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 acagagtgtttcaagtgccattcacctatgtgacaagaagaaaatggaattatctctgaa 1260

Qy       1261 catacctgtaaatcatggtccacaagaggaatcatgtggatcctctcagctccatgaaaa 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 catacctgtaaatcatggtccacaagaggaatcatgtggatcctctcagctccatgaaaa 1320

Qy       1321 tagtggttctcctgaaacttcaaggtccctgccagctcctcaagacaatgatttttttatc 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 tagtggttctcctgaaacttcaaggtccctgccagctcctcaagacaatgatttttttatc 1380

Qy       1381 tagaaaagctcaagactgttattttatgaagctgcatcactgtcctggaaatcacagttg 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 tagaaaagctcaagactgttattttatgaagctgcatcactgtcctggaaatcacagttg 1440

Qy       1441 ggatagcaccatttctggatctcaaagggctgcattctgtgatcacaagaccactccatg 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ggatagcaccatttctggatctcaaagggctgcattctgtgatcacaagaccactccatg 1500

Qy       1501 ctcttcagcaataataaatccactctcaactgcaggaaactcagaacgtctgcagcctgg 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ctcttcagcaataataaatccactctcaactgcaggaaactcagaacgtctgcagcctgg 1560

Qy       1561 tatagcccagcagtggatccagagcaaaagggaagacattgtgaaccaaatgacagaagc 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 tatagcccagcagtggatccagagcaaaagggaagacattgtgaaccaaatgacagaagc 1620

Qy       1621 ctgccttaaccagtcgctagatgcccttctgtccagggacttgatcatgaaagaggacta 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 ctgccttaaccagtcgctagatgcccttctgtccagggacttgatcatgaaagaggacta 1680

Qy       1681 tgaacttgttagtaccaagcctacaaggacctcaaaagtcagacaattactagacactac 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 tgaacttgttagtaccaagcctacaaggacctcaaaagtcagacaattactagacactac 1740

Qy       1741 tgacatccaaggagaagaatttgccaaagttatagtacaaaaattgaaagataacaaaca 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 tgacatccaaggagaagaatttgccaaagttatagtacaaaaattgaaagataacaaaca 1800

Qy       1801 aatgggtcttcagccttacccggaaatacttgtggtttctagatcaccatctttaaattt 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 aatgggtcttcagccttacccggaaatacttgtggtttctagatcaccatctttaaattt 1860

Qy       1861 acttcaaaataaaagcatgtaagtgactgtttttcaagaagaaatgtgtttcataaaagg 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1861 acttcaaaataaaagcatgtaagtgactgtttttcaagaagaaatgtgtttcataaaagg 1920

Qy   1921 atatttatatctctgttgctttgactttttttatataaaatccgtgagtattaaagcttw 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 atatttatatctctgttgctttgactttttttatataaaatccgtgagtattaaagcttw 1980

Qy   1981 awwraargktctttsrktaaatattagtctccctccatgacactgcagtatttttttaa 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 awwraargktctttsrktaaatattagtctccctccatgacactgcagtatttttttaa 2040

Qy   2041 ttaatacaagtaaaaagttgaatttgaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2098
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 ttaatacaagtaaaaagttgaatttgaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2098


RESULT   2
C77779
ID   C77779 standard; cDNA; 2709 BP.
XX
AC   C77779;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human cancer associated gene sequence SEQ ID NO:173.
XX
KW   Human; cancer associated gene; cancer antigen; detection; cancer;
KW   diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW   antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW   antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW   dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW   vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW   immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW   allergic reaction; graft versus host disease; organ rejection;
KW   haemostatic; thrombolytic; cardiovascular disorder; infection;
KW   neurological disease; drug screening; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055350-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05882.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587533/55.
DR   P-PSDB; B43570.
XX
PT   Novel isolated nucleic acids comprising sequences encoding peptides
PT   useful for treating or diagnosing e.g. cancer -
XX
PS   Claim 1; Page 751-752; 2352pp; English.
XX
CC   C77607 to C78448 encode the human cancer associated proteins given in
CC   B43398 to B44239. The proteins can have activities based on the tissues
CC   and cells the genes are expressed in. Example of activities include:
CC   cytostatic; proliferative; vulnerary; immunomodulator; antidiabetic;
CC   antiasthmatic; antirheumatic; antiarthritic; antiinflammatory;
CC   antithyroid; antiallergic; antibacterial; antiviral; dermatological;
CC   neuroprotective; cardiant; thrombolytic; coagulant; nootropic;
CC   vasotropic; antipsoriatic and antiangiogenic. The polynucleotides and
CC   polypeptides can be used for preventing, treating or ameliorating medical
CC   conditions and diagnosing pathological conditions. Polynucleotides,
CC   polypeptides, antibodies, agonists and antagonists from the present
CC   invention may be used to treat immune disorders by activating or
CC   inhibiting the proliferation, differentiation or mobilisation of immune
CC   cells, to treat disorders of haematopoietic cells, autoimmune disorders,
CC   allergic reactions, graft versus host disease and organ rejection,
```

```
CC   modulate haemostatic or thrombolytic activity, modulate inflammation,
CC   cancers, cardiovascular disorders, neurological disease and bacterial or
CC   viral infections. The peptides, nucleotides, antibodies, agonists and
CC   antagonists may be also be used in drug screens. C78449 to C78457 and
CC   B44240 represent sequences used in the exemplification of the present
CC   invention.
XX
SQ   Sequence 2709 BP; 810 A; 580 C; 540 G; 769 T; 10 other;

  Query Match             96.8%;  Score 2031.6;  DB 21;  Length 2709;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches 2028;  Conservative   12;  Mismatches    6;  Indels    0;  Gaps    0;

Qy     20 gcggcgctacggcgttggcaccagtctctagaaaagaagtcagctctggttcggagaagc 79
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     27 gcggcgctacggcgttggcaccagtctctagaaaagaagtcagctctggttcggagaagc 86

Qy     80 agcggctggcgtgggccatccggggaatgggcgccctcgtgacctagtgttgcggggcaa 139
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     87 agcggctggcgtgggccatccggggaatgggcgccctcgtgacctagtgttgcggggcaa 146

Qy    140 aaagggtcttgccggcctcgctcgtgcaggggcgtatctgggcgcctgagcgcggcgtgg 199
          |||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db    147 aaagggtcttgccggcctcgctcgtgcaggggcgtatctgggcgcctgarcgcggcgtgg 206

Qy    200 gagccttgggagccgccgcagcaggggcacacccggaaccggcctgagcgcccgggacc 259
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    207 gagccttgggagccgccgcagcaggggcacacccggaaccggcctgagcgcccgggacc 266

Qy    260 atgaacggggaggccatctgcagcgccctgcccaccattccctaccacaaactcgccgac 319
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    267 atgaacggggaggccatctgcagcgccctgcccaccattccctaccacaaactcgccgac 326

Qy    320 ctgcgctacctgagccgcggcgcctctggcactgtgtcgtccgcccgccacgcagactgg 379
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    327 ctgcgctacctgagccgcggcgcctctggcactgtgtcgtccgcccgccacgcagactgg 386

Qy    380 cgcgtccaggtggccgtgaagcacctgcacatccacactccgctgctcgacagtgaaaga 439
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    387 cgcgtccaggtggccgtgaagcacctgcacatccacactccgctgctcgacagtgaaaga 446

Qy    440 aaggatgttttaagagaagctgaaattttacacaaagctagatttagttacattttcca 499
          |||||||| ||||||||||||||||||||||||||||||||||||||||||||| |||||
Db    447 aaggatgtcttaagagaagctgaaattttacacaaagctagatttagttacattcttcca 506

Qy    500 atttgggaatttgcaatgagcctgaatttttgggaatagttactgaatacatgccaaat 559
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    507 atttgggaatttgcaatgagcctgaatttttgggaatagttactgaatacatgccaaat 566

Qy    560 ggatcattaaatgaactcctacataggaaaactgaatatcctgatgttgcttggccattg 619
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    567 ggatcattaaatgaactcctacataggaaaactgaatatcctgatgttgcttggccattg 626

Qy    620 agatttcgcatcctgcatgaaattgcccttggtgtaaattacctgcacaatatgactcct 679
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    627 agatttcgcatcctgcatgaaattgcccttggtgtaaattacctgcacaatatgactcct 686

Qy    680 cctttacttcatcatgacttgaagactcagaatatcttattggacaatgaatttcatgtt 739
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    687 cctttacttcatcatgacttgaagactcagaatatcttattggacaatgaatttcatgtt 746

Qy    740 aagattgcagattttggtttatcaaagtggcgcatgatgtccctctcacagtcacgaagt 799
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    747 aagattgcagattttggtttatcaaagtggcgcatgatgtccctctcacagtcacgaagt 806

Qy    800 agcaaatctgcaccagaaggagggacaattatttatatgccacctgaaaactatgaacct 859
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db    807 agcaaatctgcaccagaaggagggacaattatctatatgccacctgaaaactatgaacct 866

Qy    860 ggacaaaaatcaagggccagtatcaagcacgatatatatagctatgcagttatcacatgg 919
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    867 ggacaaaaatcaagggccagtatcaagcacgatatatatagctatgcagttatcacatgg 926
```

```
Qy      920 gaagtgttatccagaaaacagccttttgaagatgtcaccaatcctttgcagataatgtat 979
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      927 gaagtgttatccagaaaacagccttttgaagatgtcaccaatcctttgcagataatgtat 986

Qy      980 agtgtgtcacaaggacatcgacctgttattaatgaagaaagtttgccatatgatatacct 1039
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      987 agtgtgtcacaaggacatcgacctgttattaatgaagaaagtttgccatatgatatacct 1046

Qy     1040 caccgagcacgtatgatctctctaatagaaagtggatgggcacaaaatccagatgaaaga 1099
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1047 caccgagcacgtatgatctctctaatagaaagtggatgggcacaaaatccagatgaaaga 1106

Qy     1100 ccatctttcttaaaatgtttaatagaacttgaaccagttttgagaacatttgaagagata 1159
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1107 ccatctttcttaaaatgtttaatagaacttgaaccagttttgagaacatttgaagagata 1166

Qy     1160 acttttcttgaagctgttattcagctaaagaaaacaaagttacagagtgtttcaagtgcc 1219
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1167 acttttcttgaagctgttattcagctaaagaaaacaaagttacagagtgtttcaagtgcc 1226

Qy     1220 attcacctatgtgacaagaagaaaatggaattatctctgaacatacctgtaaatcatggt 1279
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1227 attcacctatgtgacaagaagaaaatggaattatctctgaacatacctgtaaatcatggt 1286

Qy     1280 ccacaagaggaatcatgtggatcctctcagctccatgaaaatagtggttctcctgaaact 1339
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1287 ccacaagaggaatcatgtggatcctctcagctccatgaaaatagtggttctcctgaaact 1346

Qy     1340 tcaaggtccctgccagctcctcaagacaatgatttttatctagaaaagctcaagactgt 1399
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1347 tcaaggtccctgccagctcctcaagacaatgatttttatctagaaaagctcaagactgt 1406

Qy     1400 tattttatgaagctgcatcactgtcctggaaatcacagttgggatagcaccatttctgga 1459
            |||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db     1407 tattttatgaagctgcatcactgtcctggaaatcacagttgggayagcaccatttctgga 1466

Qy     1460 tctcaaagggctgcattctgtgatcacaagaccactccatgctcttcagcaataataaat 1519
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1467 tctcaaagggctgcattctgtgatcacaagaccactccatgctcttcagcaataataaat 1526

Qy     1520 ccactctcaactgcaggaaactcagaacgtctgcagcctggtatagcccagcagtggatc 1579
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1527 ccactctcaactgcaggaaactcagaacgtctgcagcctggtatagcccagcagtggatc 1586

Qy     1580 cagagcaaaagggaagacattgtgaaccaaatgacagaagcctgccttaaccagtcgcta 1639
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1587 cagagcaaaagggaagacattgtgaaccaaatgacagaagcctgccttaaccagtcgcta 1646

Qy     1640 gatgcccttctgtccagggacttgatcatgaaagaggactatgaacttgttagtaccaag 1699
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1647 gatgcccttctgtccagggacttgatcatgaaagaggactatgaacttgttagtaccaag 1706

Qy     1700 cctacaaggacctcaaaagtcagacaattactagacactactgacatccaaggagaagaa 1759
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1707 cctacaaggacctcaaaagtcagacaattactagacactactgacatccaaggagaagaa 1766

Qy     1760 tttgccaaagttatagtacaaaaattgaaagataacaaacaaatgggtcttcagccttac 1819
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1767 tttgccaaagttatagtacaaaaattgaaagataacaaacaaatgggtcttcagccttac 1826

Qy     1820 ccggaaatacttgtggtttctagatcaccatctttaaatttacttcaaaataaaagcatg 1879
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1827 ccggaaatacttgtggtttctagatcaccatctttaaatttacttcaaaataaaagcatg 1886

Qy     1880 taagtgactgttttcaagaagaaatgtgtttcataaaaggatatttatatctctgttgc 1939
            |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db     1887 taagtgactgttttcaagaagaaatgtgtktcataaaaggatatttatatctctgttgc 1946

Qy     1940 tttgactttttttatataaaatccgtgagtattaaagcttwawwraargktctttsrkta 1999
            ||||||||||||||||||||||||||||||||||||||||:|:::||:|:|||||:::||
Db     1947 tttgactttttttatataaaatccgtgagtattaaagctttattgaaggttctttgggta 2006

Qy     2000 aatattagtctccctccatgacactgcagtattttttttaattaatacaagtaaaaagtt 2059
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2007 aatattagtctccctccatgacactgcagtattttttttaattaatacaagtaaaaagtt 2066

Qy     2060 gaattt 2065
              | ||
Db     2067 tgaatt 2072


RESULT   3
Z48762
ID   Z48762 standard; cDNA; 2502 BP.
XX
AC   Z48762;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Human RICK coding sequence.
XX
KW   RICK; human; RIP-like interacting CLARP kinase; apoptosis regulator; ARC;
KW   caspase-8; caspase-10; CD95 signalling; apoptosis signalling pathway;
KW   CIDE-A; CIDE-B; DREP-1; diagnosis; cell growth; apoptosis dysregulation;
KW   gene therapy; cardiac disorder; cancer; neurodegenerative disorder; AIDS;
KW   aplastic anaemia; ischaemic injury; toxin-induced liver disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO9955134-A2.
XX
PD   04-NOV-1999.
XX
PF   27-APR-1999;   99WO-US09183.
XX
PR   27-APR-1998;   98US-0069023.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Nunez G,  Inohara N,  Koseki T;
XX
DR   WPI; 2000-072163/06.
DR   P-PSDB; Y59404.
XX
PT   Compositions for identifying apoptosis signalling pathway inhibitors
PT   useful for treating diseases -
XX
PS   Claim 8; Fig 7b; 93pp; English.
XX
CC   This sequence encodes the human RICK (RIP-like interacting CLARP kinase)
CC   protein of the invention. The RICK protein acts as a positive regulator
CC   of apoptosis, potentiating apoptosis induced by caspase-8 and caspase-10
CC   during CD95 signalling. The invention provides methods for identifying
CC   apoptosis signalling pathway inhibitors and activators, and methods and
CC   compositions for screening compounds which will modulate the interactions
CC   of the various compositions identified: ARC, RICK, and the CIDE family of
CC   activators (CIDE-A, CIDE-B and DREP-1). RICK is useful in screening
CC   assays for agents, useful in the diagnosis, prognosis or treatment of
CC   disease associated with excess cell growth and dysregulation of
CC   apoptosis. Complexes containing RICK and CLARP can be used in drug
CC   screening assays to identify inhibitor molecules blocking CD95-mediated
CC   apoptosis. Overexpression of ARC in an in vitro cell system can be used
CC   to identify inhibitors of the enzymatic activity of caspase-8.
CC   Identification of ARC-like inhibitory compounds may be useful for gene
CC   therapy treatment of disease with increased cell death in muscle tissue
CC   and cardiac disorders. Therapeutic compositions of CIDEs can be used to
CC   treat e.g. cancer, AIDS, neurodegenerative disorders, aplastic anaemia,
CC   ischaemic injury, and toxin-induced liver disease. AntiRICK antibodies
CC   can be used as reagents for the preparation or affinity chromatography
CC   media, and for diagnostically measuring RICK levels. A specific inhibitor
CC   of an essential step in the biochemistry of apoptosis is needed. RICK
CC   interaction with intracellular factors such as CLARP and FADD appears to
CC   be essential for apoptosis, inhibitors of RICK binding to intracellular
CC   apoptosis factors are potential drug candidates.
XX
```

```
SQ   Sequence 2502 BP; 769 A; 535 C; 499 G; 699 T; 0 other;

  Query Match             95.7%;  Score 2006.8;  DB 21;  Length 2502;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 2016;  Conservative    9;  Mismatches    6;  Indels    1;  Gaps    1;

Qy     36 ggcaccagtctctagaaaagaagtcagctctggttcggagaagcagcggctggcgtgggc 95
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ggcaccagtctctagaaaagaagtcagctctggttcggagaagcagcggctggcgtgggc 60

Qy     96 catccggggaatgggcgccctcgtgacctagtgttgcggggcaaaaagggtcttgccggc 155
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 catccggggaatgggcgccctcgtgacctagtgttgcggggcaaaaagggtcttgccggc 120

Qy    156 ctcgctcgtgcaggggcgtatctgggcgcctgagcgcggcgtgggagccttgggagccgc 215
          ||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db    121 ctcgctcgtgcaggggcgtatctgggcgcctgagcgcgcagtgggagccttgggagccgc 180

Qy    216 cgcagcagggggcacacccggaaccggcctgagcgcccgggaccatgaacggggaggcca 275
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 cgcagcagggggcacacccggaaccggcctgagcgcccgggaccatgaacggggaggcca 240

Qy    276 tctgcagcgccc-tgcccaccattccctaccacaaactcgccgacctgcgctacctgagc 334
          |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db    241 tctgcagcgcccatgcccaccattccctaccacaaactcgccgacctgcgctacctgagc 300

Qy    335 cgcggcgcctctggcactgtgtcgtccgcccgccacgcagactggcgcgtccaggtggcc 394
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 cgcggcgcctctggcactgtgtcgtccgcccgccacgcagactggcgcgtccaggtggcc 360

Qy    395 gtgaagcacctgcacatccacactccgctgctcgacagtgaaagaaaggatgttttaaga 454
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    361 gtgaagcacctgcacatccacactccgctgctcgacagtgaaagaaaggatgtcttaaga 420

Qy    455 gaagctgaaattttacacaaagctagatttagttacatttttccaattttgggaatttgc 514
          ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    421 gaagctgaaattttacacaaagctagatttagttacattcttccaattttgggaatttgc 480

Qy    515 aatgagcctgaatttttgggaatagttactgaatacatgccaaatggatcattaaatgaa 574
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 aatgagcctgaatttttgggaatagttactgaatacatgccaaatggatcattaaatgaa 540

Qy    575 ctcctacataggaaaactgaatatcctgatgttgcttggccattgagatttcgcatcctg 634
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ctcctacataggaaaactgaatatcctgatgttgcttggccattgagatttcgcatcctg 600

Qy    635 catgaaattgcccttggtgtaaattacctgcacaatatgactcctcctttacttcatcat 694
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 catgaaattgcccttggtgtaaattacctgcacaatatgactcctcctttacttcatcat 660

Qy    695 gacttgaagactcagaatatcttattggacaatgaatttcatgttaagattgcagatttt 754
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 gacttgaagactcagaatatcttattggacaatgaatttcatgttaagattgcagatttt 720

Qy    755 ggtttatcaaagtggcgcatgatgtccctctcacagtcacgaagtagcaaatctgcacca 814
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 ggtttatcaaagtggcgcatgatgtccctctcacagtcacgaagtagcaaatctgcacca 780

Qy    815 gaaggagggacaattatttatatgccacctgaaaactatgaacctggacaaaaatcaagg 874
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    781 gaaggagggacaattatctatatgccacctgaaaactatgaacctggacaaaaatcaagg 840

Qy    875 gccagtatcaagcacgatatatatagctatgcagttatcacatgggaagtgttatccaga 934
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 gccagtatcaagcacgatatatatagctatgcagttatcacatgggaagtgttatccaga 900

Qy    935 aaacagccttttgaagatgtcaccaatcctttgcagataatgtatagtgtgtcacaagga 994
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 aaacagccttttgaagatgtcaccaatcctttgcagataatgtatagtgtgtcacaagga 960

Qy    995 catcgacctgttattaatgaagaaagtttgccatatgatatacctcaccgagcacgtatg 1054
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 catcgacctgttattaatgaagaaagtttgccatatgatatacctcaccgagcacgtatg 1020

Qy   1055 atctctctaatagaaagtggatgggcacaaaatccagatgaaagaccatctttcttaaaa 1114
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 atctctctaatagaaagtggatgggcacaaaatccagatgaaagaccatctttcttaaaa 1080

Qy   1115 tgtttaatagaacttgaaccagttttgagaacatttgaagagataacttttcttgaagct 1174
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 tgtttaatagaacttgaaccagttttgagaacatttgaagagataacttttcttgaagct 1140

Qy   1175 gttattcagctaaagaaaacaaagttacagagtgtttcaagtgccattcacctatgtgac 1234
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 gttattcagctaaagaaaacaaagttacagagtgtttcaagtgccattcacctatgtgac 1200

Qy   1235 aagaagaaaatggaattatctctgaacatacctgtaaatcatggtccacaagaggaatca 1294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 aagaagaaaatggaattatctctgaacatacctgtaaatcatggtccacaagaggaatca 1260

Qy   1295 tgtggatcctctcagctccatgaaaatagtggttctcctgaaacttcaaggtccctgcca 1354
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 tgtggatcctctcagctccatgaaaatagtggttctcctgaaacttcaaggtccctgcca 1320

Qy   1355 gctcctcaagacaatgattttttatctagaaaagctcaagactgttattttatgaagctg 1414
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 gctcctcaagacaatgattttttatctagaaaagctcaagactgttattttatgaagctg 1380

Qy   1415 catcactgtcctggaaatcacagttgggatagcaccatttctggatctcaaagggctgca 1474
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   1381 catcactgtcctggaaatcacagttgggatagcaccatttctggttctcaaagggctgca 1440

Qy   1475 ttctgtgatcacaagaccactccatgctcttcagcaataataaatccactctcaactgca 1534
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 ttctgtgatcacaagaccactccatgctcttcagcaataataaatccactctcaactgca 1500

Qy   1535 ggaaactcagaacgtctgcagcctggtatagcccagcagtggatccagagcaaaagggaa 1594
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ggaaactcagaacgtctgcagcctggtatagcccagcagtggatccagagcaaaagggaa 1560

Qy   1595 gacattgtgaaccaaatgacagaagcctgccttaaccagtcgctagatgcccttctgtcc 1654
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 gacattgtgaaccaaatgacagaagcctgccttaaccagtcgctagatgcccttctgtcc 1620

Qy   1655 agggacttgatcatgaaagaggactatgaacttgttagtaccaagcctacaaggacctca 1714
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 agggacttgatcatgaaagaggactatgaacttgttagtaccaagcctacaaggacctca 1680

Qy   1715 aaagtcagacaattactagacactactgacatccaaggagaagaatttgccaaagttata 1774
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 aaagtcagacaattactagacactactgacatccaaggagaagaatttgccaaagttata 1740

Qy   1775 gtacaaaaattgaaagataacaaacaaatgggtcttcagccttacccggaaatacttgtg 1834
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 gtacaaaaattgaaagataacaaacaaatgggtcttcagccttacccggaaatacttgtg 1800

Qy   1835 gtttctagatcaccatctttaaatttacttcaaaataaaagcatgtaagtgactgttttt 1894
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 gtttctagatcaccatctttaaatttacttcaaaataaaagcatgtaagtgactgttttt 1860

Qy   1895 caagaagaaatgtgtttcataaaaggatatttatatctctgttgctttgactttttttat 1954
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 caagaagaaatgtgtttcataaaaggatatttatatctctgttgctttgactttttttat 1920

Qy   1955 ataaaatccgtgagtattaaagcttwawwraargktctttsrktaaatattagtctccct 2014
          |||||||||||||||||||||||||:|:::||:|:|||||:::|||||||||||||||||
Db   1921 ataaaatccgtgagtattaaagctttattgaaggttctttgggtaaatattagtctccct 1980

Qy   2015 ccatgacactgcagtattttttttaattaatacaagtaaaaagttgaatttg 2066
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ccatgacactgcagtattttttttaattaatacaagtaaaaagttgaatttg 2032

RESULT   4
Z46143
```

```
ID   Z46143 standard; cDNA; 2024 BP.
XX
AC   Z46143;
XX
DT   16-MAY-2000  (first entry)
XX
DE   cDNA sequence encoding a human phosphorylation effector PHSP-6.
XX
KW   Human; phosphorylation effector; PHSP; proliferative disorder;
KW   immune disorder; neuronal disorder; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             203..1825
FT                   /*tag=  a
FT                   /product= "phosphorylation effector"
XX
PN   WO200006728-A2.
XX
PD   10-FEB-2000.
XX
PF   28-JUL-1999;   99WO-US17132.
XX
PR   28-JUL-1998;   98US-0123494.
PR   14-SEP-1998;   98US-0152814.
PR   14-OCT-1998;   98US-0173482.
PR   03-NOV-1998;   98US-0106889.
PR   19-NOV-1998;   98US-0109093.
PR   22-DEC-1998;   98US-0113796.
PR   12-JAN-1999;   99US-0173482.
PR   12-JAN-1999;   99US-0229005.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Hillman JL,  Lal P,  Tang YT,  Corley NC,  Guegler KJ,  Baughn MR;
PI   Patterson C,  Bandman O,  Au-Young J,  Gorgone GA,  Yue H,  Azimzai Y;
PI   Reddy R,  Lu DAM,  Shih LL;
XX
DR   WPI; 2000-183125/16.
DR   P-PSDB; Y68774.
XX
PT   New human phosphorylation effectors useful for the diagnosis, treatment
PT   and prevention of proliferative, immune and neuronal disorders  -
XX
PS   Claim 9; Page 121-122; 142pp; English.
XX
CC   Z46138-Z46168 encode human phosphorylation effectors (PHSP),
CC   designated PHSP1-PHSP31 (the protein sequence for PHSP28 is not given
CC   in the specification). The sequences were isolated from cDNA libraries
CC   prepared from various human tissues. The PHSP proteins are useful for
CC   the diagnosis, treatment and prevention of proliferative disorders,
CC   immune disorders and neuronal disorders. The PHSP proteins form
CC   pharmaceutical compositions which useful for treating or preventing
CC   disorders associated with decreased PHSP expression/activity. PHSP
CC   antagonists are useful for treating or preventing disorders associated
CC   with increased PHSP expression/activity.
XX
SQ   Sequence 2024 BP; 612 A; 445 C; 434 G; 533 T; 0 other;


  Query Match             95.3%;  Score 1999.6;  DB 21;  Length 2024;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 2002;  Conservative    9;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         58 gtcagctctggttcggagaagcagcggctggcgtgggccatccggggaatgggcgccctc 117
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 gtcagctctggttcggagaagcagcggctggcgtgggccatccggggaatgggcgccctc 60

Qy        118 gtgacctagtgttgcggggcaaaaagggtcttgccggcctcgctcgtgcaggggcgtatc 177
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 gtgacctagtgttgcggggcaaaaagggtcttgccggcctcgctcgtgcaggggcgtatc 120

Qy        178 tgggcgcctgagcgcggcgtgggagccttgggagccgccgcagcagggggcacacccgga 237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 tgggcgcctgagcgcggcgtgggagccttgggagccgccgcagcagggggcacacccgga 180

Qy        238 accggcctgagcgcccgggaccatgaacggggaggccatctgcagcgccctgcccaccat 297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 accggcctgagcgcccgggaccatgaacggggaggccatctgcagcgccctgcccaccat 240

Qy        298 tccctaccacaaactcgccgacctgcgctacctgagccgcggcgcctctggcactgtgtc 357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 tccctaccacaaactcgccgacctgcgctacctgagccgcggcgcctctggcactgtgtc 300

Qy        358 gtccgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctgcacatccacac 417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 gtccgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctgcacatccacac 360

Qy        418 tccgctgctcgacagtgaaagaaaggatgttttaagagaagctgaaatttacacaaagc 477
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        361 tccgctgctcgacagtgaaagaaaggatgtcttaagagaagctgaaatttacacaaagc 420

Qy        478 tagatttagttacatttttccaattttgggaatttgcaatgagcctgaatttttgggaat 537
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        421 tagatttagttacattcttccaattttgggaatttgcaatgagcctgaatttttgggaat 480

Qy        538 agttactgaatacatgccaaatggatcattaaatgaactcctacataggaaaactgaata 597
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 agttactgaatacatgccaaatggatcattaaatgaactcctacataggaaaactgaata 540

Qy        598 tcctgatgttgcttggccattgagatttcgcatcctgcatgaaattgcccttggtgtaaa 657
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 tcctgatgttgcttggccattgagatttcgcatcctgcatgaaattgcccttggtgtaaa 600

Qy        658 ttacctgcacaatatgactcctcctttacttcatcatgacttgaagactcagaatatctt 717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ttacctgcacaatatgactcctcctttacttcatcatgacttgaagactcagaatatctt 660

Qy        718 attggacaatgaatttcatgttaagattgcagattttggtttatcaaagtggcgcatgat 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 attggacaatgaatttcatgttaagattgcagattttggtttatcaaagtggcgcatgat 720

Qy        778 gtccctctcacagtcacgaagtagcaaatctgcaccagaaggagggacaattatttatat 837
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        721 gtccctctcacagtcacgaagtagcaaatctgcaccagaaggagggacaattatctatat 780

Qy        838 gccacctgaaaactatgaacctggacaaaaatcaagggccagtatcaagcacgatatata 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 gccacctgaaaactatgaacctggacaaaaatcaagggccagtatcaagcacgatatata 840

Qy        898 tagctatgcagttatcacatgggaagtgttatccagaaaacagccttttgaagatgtcac 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 tagctatgcagttatcacatgggaagtgttatccagaaaacagccttttgaagatgtcac 900

Qy        958 caatcctttgcagataatgtatagtgtgtcacaaggacatcgacctgttattaatgaaga 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 caatcctttgcagataatgtatagtgtgtcacaaggacatcgacctgttattaatgaaga 960

Qy       1018 aagtttgccatatgatatacctcaccgagcacgtatgatctctctaatagaaagtggatg 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 aagtttgccatatgatatacctcaccgagcacgtatgatctctctaatagaaagtggatg 1020

Qy       1078 ggcacaaaatccagatgaaagaccatctttcttaaaatgtttaatagaacttgaaccagt 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ggcacaaaatccagatgaaagaccatctttcttaaaatgtttaatagaacttgaaccagt 1080

Qy       1138 tttgagaacatttgaagagataacttttcttgaagctgttattcagctaaagaaaacaaa 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 tttgagaacatttgaagagataacttttcttgaagctgttattcagctaaagaaaacaaa 1140

Qy       1198 gttacagagtgtttcaagtgccattcacctatgtgacaagaagaaaatggaattatctct 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 gttacagagtgtttcaagtgccattcacctatgtgacaagaagaaaatggaattatctct 1200

Qy       1258 gaacatacctgtaaatcatggtccacaagaggaatcatgtggatcctctcagctccatga 1317
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 gaacatacctgtaaatcatggtccacaagaggaatcatgtggatcctctcagctccatga 1260

Qy     1318 aaatagtggttctcctgaaacttcaaggtccctgccagctcctcaagacaatgatttttt 1377
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 aaatagtggttctcctgaaacttcaaggtccctgccagctcctcaagacaatgatttttt 1320

Qy     1378 atctagaaaagctcaagactgttattttatgaagctgcatcactgtcctggaaatcacag 1437
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 atctagaaaagctcaagactgttattttatgaagctgcatcactgtcctggaaatcacag 1380

Qy     1438 ttgggatagcaccatttctggatctcaaagggctgcattctgtgatcacaagaccactcc 1497
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 ttgggatagcaccatttctggatctcaaagggctgcattctgtgatcacaagaccactcc 1440

Qy     1498 atgctcttcagcaataataaatccactctcaactgcaggaaactcagaacgtctgcagcc 1557
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1441 atgctcttcagcaataataaatccactctcaactgcaggaaactcagaacgtctgcagcc 1500

Qy     1558 tggtatagcccagcagtggatccagagcaaaagggaagacattgtgaaccaaatgacaga 1617
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1501 tggtatagcccagcagtggatccagagcaaaagggaagacattgtgaaccaaatgacaga 1560

Qy     1618 agcctgccttaaccagtcgctagatgcccttctgtccagggacttgatcatgaaagagga 1677
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1561 agcctgccttaaccagtcgctagatgcccttctgtccagggacttgatcatgaaagagga 1620

Qy     1678 ctatgaacttgttagtaccaagcctacaaggacctcaaaagtcagacaattactagacac 1737
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1621 ctatgaacttgttagtaccaagcctacaaggacctcaaaagtcagacaattactagacac 1680

Qy     1738 tactgacatccaaggagaagaatttgccaaagttatagtacaaaaattgaaagataacaa 1797
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1681 tactgacatccaaggagaagaatttgccaaagttatagtacaaaaattgaaagataacaa 1740

Qy     1798 acaaatgggtcttcagccttacccggaaatacttgtggtttctagatcaccatctttaaa 1857
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1741 acaaatgggtcttcagccttacccggaaatacttgtggtttctagatcaccatctttaaa 1800

Qy     1858 tttacttcaaaataaaagcatgtaagtgactgtttttcaagaagaaatgtgtttcataaa 1917
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1801 tttacttcaaaataaaagcatgtaagtgactgtttttcaagaagaaatgtgtttcataaa 1860

Qy     1918 aggatatttatatctctgttgctttgactttttttatataaaatccgtgagtattaaagc 1977
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1861 aggatatttatatctctgttgctttgactttttttatataaaatccgtgagtattaaagc 1920

Qy     1978 ttwawwraargktctttsrktaaatattagtctccctccatgacactgcagtatttttt 2037
               ||:|:::||:|:|||||:::||||||||||||||||||||||||||||||||||||||||
Db     1921 tttattgaaggttctttgggtaaatattagtctccctccatgacactgcagtatttttt 1980

Qy     2038 taattaatacaagtaaaaagttgaatttgaaaaaaaaaaaaaaa 2081
               ||||||||||||||||||||||  |  ||     |  | |||||||
Db     1981 taattaatacaagtaaaaagtttgaattttgctacataaaaaaa 2024

RESULT    5
Z09246
ID   Z09246 standard; cDNA; 1931 BP.
XX
AC   Z09246;
XX
DT   25-OCT-1999  (first entry)
XX
DE   Human CARD-3 cDNA.
XX
KW   CARD-3; caspase recruitment domain; CARD-4; regulation; detection;
KW   caspase activation; detection; screening; therapy; diagnosis; disease;
KW   apoptotic cell death; Fas/APO-1 receptor complex; TNF receptor complex;
KW   cancer; follicular lymphoma; carcinoma; p53 mutation; viral infection;
KW   hormone-dependent tumour; autoimmune disorder; Alzheimer's disease;
KW   systemic lupus erythematosis; immune-mediated glomerulonephritis; stroke;
KW   Parkinson's disease; amyotrophic lateral sclerosis; retinitis pigmentosa;
```

```
KW   spinal muscular dystrophy; cerebellar degeneration; anaemia; drug;
KW   myelodysplastic syndrome; myocardial infarction; cell proliferation;
KW   cell differentiation; cell survival; CARD-4L; CARD-4S; CARD-4Y;
KW   CARD-4Z; human; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             214..1836
FT                   /*tag=  a
FT                   /product= "CARD-3"
XX
PN   WO9940102-A1.
XX
PD   12-AUG-1999.
XX
PF   05-FEB-1999;   99WO-US02544.
XX
PR   08-DEC-1998;   98US-0207359.
PR   06-FEB-1998;   98US-0019942.
PR   17-JUN-1998;   98US-0099041.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Bertin J;
XX
DR   WPI; 1999-494269/41.
DR   P-PSDB; Y31140.
XX
PT   Novel CARD-3 and CARD-4 genes and polypeptides used or treating
PT   regulation of cellular proliferation and differentiation and cell
PT   survival
XX
PS   Example 2; Fig 1; 181pp; English.
XX
CC   This invention describes the isolation of novel human caspase
CC   recruitment domain, CARD-3 and CARD-4 polynucleotides and proteins and a
CC   partial murine CARD-4L protein and genes. The genes and proteins of
CC   the invention are involved in the regulation of caspase activation.
CC   The caspase recruitment domain (CARD) polynucleotides, polypeptides,
CC   homologues and antibodies can be used in screening assays, detection
CC   assays, predictive medicine and therapeutic and prophylactic methods of
CC   treatment. The methods may be used to diagnose and treat patients which
CC   are suffering from a disorder associated with abnormal level or rate of
CC   apoptotic cell death, abnormal activity of the Fas/APO-1 receptor
CC   complex, abnormal activity of the TNF receptor complex, or abnormal
CC   activity of a caspase. Diseases that may be treated include cancer
CC   (particularly follicular lymphoma, carcinomas associated with mutations
CC   in p53 and hormone-dependent tumours), autoimmune disorders (e.g.
CC   systemic lupus erythematosis, immune-mediated glomerulonephritis), viral
CC   infections, Alzheimer's disease, Parkinson's disease, amyotrophic lateral
CC   sclerosis, retinitis pigmentosa, spinal muscular dystrophy, cerebellar
CC   degeneration, anaemia, myelodysplastic syndrome, myocardial infarction,
CC   and stroke. CARD-3 protein interacts with other cellular proteins, and so
CC   can be used for regulation of cellular proliferation and differentiation
CC   and cell survival. The CARD proteins may also be used to for screen drugs
CC   or compounds which modulate their activity. The CARD-4 gene can express a
CC   long transcript that encodes CARD-4L, a short transcript that encodes
CC   CARD-4S or two CARD-4 splice variants, CARD-4Y and CARD-4Z. This sequence
CC   encodes the human CARD-3 protein described in the method of the
CC   invention.
XX
SQ   Sequence 1931 BP; 613 A; 429 C; 416 G; 473 T; 0 other;

  Query Match             88.8%;  Score 1864;  DB 20;  Length 1931;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1867;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       58 gtcagctctggttcggagaagcagcggctggcgtgggccatccggggaatgggcgccctc 117
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       12 gtcagctctggttcggagaagcagcggctggcgtgggccatccggggaatgggcgccctc 71
```

```
Qy    118 gtgacctagtgttgcggggcaaaaagggtcttgccggcctcgctcgtgcaggggcgtatc 177
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     72 gtgacctagtgttgcggggcaaaaagggtcttgccggcctcgctcgtgcaggggcgtatc 131

Qy    178 tgggcgcctgagcgcggcgtgggagccttgggagccgccgcagcaggggcacacccgga 237
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    132 tgggcgcctgagcgcggcgtgggagccttgggagccgccgcagcaggggcacacccgga 191

Qy    238 accggcctgagcgcccgggaccatgaacggggaggccatctgcagcgccctgcccaccat 297
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    192 accggcctgagcgcccgggaccatgaacggggaggccatctgcagcgccctgcccaccat 251

Qy    298 tccctaccacaaactcgccgacctgcgctacctgagccgcggcgcctctggcactgtgtc 357
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    252 tccctaccacaaactcgccgacctgcgctacctgagccgcggcgcctctggcactgtgtc 311

Qy    358 gtccgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctgcacatccacac 417
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    312 gtccgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctgcacatccacac 371

Qy    418 tccgctgctcgacagtgaaagaaaggatgttttaagagaagctgaaattttacacaaagc 477
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db    372 tccgctgctcgacagtgaaagaaaggatgtcttaagagaagctgaaattttacacaaagc 431

Qy    478 tagatttagttacattttttccaatttgggaatttgcaatgagcctgaatttttgggaat 537
          |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db    432 tagatttagttacattcttccaatttgggaatttgcaatgagcctgaatttttgggaat 491

Qy    538 agttactgaatacatgccaaatggatcattaaatgaactcctacataggaaaactgaata 597
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    492 agttactgaatacatgccaaatggatcattaaatgaactcctacataggaaaactgaata 551

Qy    598 tcctgatgttgcttggccattgagatttcgcatcctgcatgaaattgcccttggtgtaaa 657
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db    552 tcctgatgttgcttggccattgagatttcgtatcctgcatgaaattgcccttggtgtaaa 611

Qy    658 ttacctgcacaatatgactcctcctttacttcatcatgacttgaagactcagaatatctt 717
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    612 ttacctgcacaatatgactcctcctttacttcatcatgacttgaagactcagaatatctt 671

Qy    718 attggacaatgaatttcatgttaagattgcagattttggtttatcaaagtggcgcatgat 777
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    672 attggacaatgaatttcatgttaagattgcagattttggtttatcaaagtggcgcatgat 731

Qy    778 gtccctctcacagtcacgaagtagcaaatctgcaccagaaggagggacaattatttatat 837
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db    732 gtccctctcacagtcacgaagtagcaaatctgcaccagaaggagggacaattatctatat 791

Qy    838 gccacctgaaaactatgaacctggacaaaaatcaagggccagtatcaagcacgatatata 897
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    792 gccacctgaaaactatgaacctggacaaaaatcaagggccagtatcaagcacgatatata 851

Qy    898 tagctatgcagttatcacatgggaagtgttatccagaaaacagccttttgaagatgtcac 957
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    852 tagctatgcagttatcacatgggaagtgttatccagaaaacagccttttgaagatgtcac 911

Qy    958 caatcctttgcagataatgtatagtgtgtcacaaggacatcgacctgttattaatgaaga 1017
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    912 caatcctttgcagataatgtatagtgtgtcacaaggacatcgacctgttattaatgaaga 971

Qy   1018 aagtttgccatatgatatacctcaccgagcacgtatgatctctctaatagaaagtggatg 1077
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    972 aagtttgccatatgatatacctcaccgagcacgtatgatctctctaatagaaagtggatg 1031

Qy   1078 ggcacaaaatccagatgaaagaccatctttcttaaaatgtttaatagaacttgaaccagt 1137
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1032 ggcacaaaatccagatgaaagaccatctttcttaaaatgtttaatagaacttgaaccagt 1091

Qy   1138 tttgagaacatttgaagagataacttttcttgaagctgttattcagctaaagaaaacaaa 1197
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1092 tttgagaacatttgaagagataacttttcttgaagctgttattcagctaaagaaaacaaa 1151

Qy   1198 gttacagagtgtttcaagtgccattcacctatgtgacaagaagaaaatggaattatctct 1257
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1152 gttacagagtgtttcaagtgccattcacctatgtgacaagaagaaaatggaattatctct 1211

Qy   1258 gaacatacctgtaaatcatggtccacaagaggaatcatgtggatcctctcagctccatga 1317
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1212 gaacatacctgtaaatcatggtccacaagaggaatcatgtggatcctctcagctccatga 1271

Qy   1318 aaatagtggttctcctgaaacttcaaggtccctgccagctcctcaagacaatgatttttt 1377
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1272 aaatagtggttctcctgaaacttcaaggtccctgccagctcctcaagacaatgatttttt 1331

Qy   1378 atctagaaaagctcaagactgttattttatgaagctgcatcactgtcctggaaatcacag 1437
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1332 atctagaaaagctcaagactgttattttatgaagctgcatcactgtcctggaaatcacag 1391

Qy   1438 ttgggatagcaccatttctggatctcaaagggctgcattctgtgatcacaagaccactcc 1497
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   1392 ttgggatagcaccatttctggatctcaaagggctgcattctgtgatcacaagaccattcc 1451

Qy   1498 atgctcttcagcaataataaatccactctcaactgcaggaaactcagaacgtctgcagcc 1557
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1452 atgctcttcagcaataataaatccactctcaactgcaggaaactcagaacgtctgcagcc 1511

Qy   1558 tggtatagcccagcagtggatccagagcaaaagggaagacattgtgaaccaaatgacaga 1617
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1512 tggtatagcccagcagtggatccagagcaaaagggaagacattgtgaaccaaatgacaga 1571

Qy   1618 agcctgccttaaccagtcgctagatgcccttctgtccagggacttgatcatgaaagagga 1677
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1572 agcctgccttaaccagtcgctagatgcccttctgtccagggacttgatcatgaaagagga 1631

Qy   1678 ctatgaacttgttagtaccaagcctacaaggacctcaaaagtcagacaattactagacac 1737
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1632 ctatgaacttgttagtaccaagcctacaaggacctcaaaagtcagacaattactagacac 1691

Qy   1738 tactgacatccaaggagaagaatttgccaaagttatagtacaaaaattgaaagataacaa 1797
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1692 tactgacatccaaggagaagaatttgccaaagttatagtacaaaaattgaaagataacaa 1751

Qy   1798 acaaatgggtcttcagccttacccggaaatacttgtggtttctagatcaccatctttaaa 1857
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1752 acaaatgggtcttcagccttacccggaaatacttgtggtttctagatcaccatctttaaa 1811

Qy   1858 tttacttcaaaataaaagcatgtaagtgactgtttttcaagaagaaatgtgtttcataaa 1917
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1812 tttacttcaaaataaaagcatgtaagtgactgtttttcaagaagaaatgtgtttcataaa 1871

Qy   1918 aggatatttata 1929
          ||||||||||||
Db   1872 aggatatttata 1883


RESULT   6
F21841
ID   F21841 standard; DNA; 463 BP.
XX
AC   F21841;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Human breast and ovarian cancer associated antigen gene SEQ ID 228.
XX
KW   Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW   nootropic; neurpprotective; antiviral; antiallergic; hepatotropic;
KW   antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;
KW   antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
KW   Addison's disease; allergy; autoimmune haemolytic anaemia;
KW   autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW   multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW   cardiovascular disorder; wound healing; neurological disease; ds.
XX
OS   Homo sapiens.
XX
```

```
PN   WO200055173-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05881.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-611515/58.
DR   P-PSDB; B58938.
XX
PT   New human breast and ovarian cancer associated gene sequences and the
PT   polypeptides encoded by these genes, useful in the prevention,
PT   treatment and diagnosis of cancer, immune disorders, cardiovascular
PT   disorders and neurological diseases -
XX
PS   Claim 1; Page 653; 1299pp; English.
XX
CC   Sequences F21614 - F22031 represent DNA sequences encoding human proteins
CC   B58711 - B59128. The DNA and protein sequences are associated with
CC   breast and ovarian cancer. Included in the invention are sequences
CC   F22032 - F22040 and B59129 which are used in the isolation and
CC   characterisation of the DNA and protein sequences of the invention. The
CC   breast and ovarian cancer associated DNA, protein, agonist or antagonist
CC   sequences exhibit cytostatic; immunosuppressive; nootropic;
CC   neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
CC   antiinflammatory; antiulcer; vulnerary; anticonvulsant; antibacterial;
CC   antifungal; antiparasitic and cardiant activity. The polynucleotide and
CC   protein sequences are used in the diagnosis of cancer, particularly
CC   breast and ovarian cancer. The nucleic acid sequences, proteins, agonists
CC   and agonists may also be used in the diagnosis, prevention and treatment
CC   of immune disorders e.g. Addison's disease, allergies, autoimmune
CC   haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC   disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC   cardiovascular disorders such as myocardial ischaemias; wound healing;
CC   neurological diseases such as cerebral anoxia and epilepsy; and
CC   infectious diseases.
XX
SQ   Sequence 463 BP; 150 A; 92 C; 89 G; 131 T; 1 other;

  Query Match             21.7%;  Score 454.6;  DB 21;  Length 463;
  Best Local Similarity   98.7%;  Pred. No. 1.6e-80;
  Matches  457;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy    546 aatacatgccaaatggatcattaaatgaactcctacataggaaaactgaatatcctgatg 605
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 aatacatgccaaatggatcattaaatgaactcctacataggaaaactgaatatcctgatg 60

Qy    606 ttgcttggccattgagatttcgcatcctgcatgaaattgcccttggtgtaaattacctgc 665
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ttgcttggccattgagatttcgcatcctgcatgaaattgcccttggtgtaaattacctgc 120

Qy    666 acaatatgactcctcctttacttcatcatgacttgaagactcagaatatcttattggaca 725
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 acaatatgactcctcctttacttcatcatgacttgaagactcagaatatcttattggaca 180

Qy    726 atgaatttcatgttaagattgcagattttggtttatcaaagtggcgcatgatgtccctct 785
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 atgaatttcatgttaagattgcagattttggtttatcaaagtggcgcatgatgtccctct 240

Qy    786 cacagtcacgaagtagcaaatctgcaccagaaggagggacaattatttatatgccacctg 845
          |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db    241 cacagtcacgaagtagcaaatctgcaccagaaggagggacaattatctatatgccacctg 300

Qy    846 aaaactatgaacctggacaaaaatcaagggccagtatcaagcacgatatatatagctatg 905
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 aaaactatgaacctggacaaaaatcaagggccagtatcaagcacgatatatatagctatg 360
```

```
Qy    906 cagttatcacatgggaagtgttatccagaaaacagccttttgaagatgtcaccaatcctt 965
          ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db    361 cagttatcacatgggaagtgktatccagaaaacagccttttgaagatgtcaccaatcctt 420

Qy    966 tgcagataatgtatagtgtgtcacaaggacatcgacctgttat 1008
          |||||||||||||||||||||||||||||||| |  ||| |||
Db    421 tgcagataatgtatagtgtgtcacaaggacattggactggtat 463


RESULT   7
V33547
ID   V33547 standard; cDNA; 210 BP.
XX
AC   V33547;
XX
DT   29-DEC-1998  (first entry)
XX
DE   Clone 156108 cDNA fragment encoding a human protein kinase homolog.
XX
KW   Protein kinase; cell signalling; inflammation; carcinoma; diabetes;
KW   human X-linked agammaglobulinaemia; nonspherocytic haemolytic anaemia;
KW   artherosclerosis; glioma; restenosis; cholera-based septic shock;
KW   U01064 Dictyostelium Y kinase; ss.
XX
OS   Homo sapiens.
XX
PN   US5817479-A.
XX
PD   06-OCT-1998.
XX
PF   07-AUG-1996;   96US-0700575.
XX
PR   07-AUG-1996;   96US-0700575.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Au-Young J,  Bandman O,  Hawkins PR,  Wilde CG;
XX
DR   WPI; 1998-556387/47.
XX
PT   Human kinase poly:nucleotide(s) and recombinant products - useful
PT   for identification of modulators of the enzyme, and treatment of
PT   diseases associated with abnormal kinase expression
XX
PS   Claim 1; Columns 45-48; 30pp; English.
XX
CC   The invention provides polynucleotides which encode novel protein
CC   kinase homologs expressed in various human cells and tissues.  The
CC   present sequence represents the clone 156108 cDNA fragment derived from
CC   a human cDNA library designated THP-1 phorbol LPS.  The cDNA encodes
CC   a protein kinase which shows homology to the U01064 Dictyostelium Y
CC   kinase.  Vectors and host cells can be used for recombinant production
CC   of the protein kinase homolog.  The recombinant proteins may be used to
CC   raise antibodies for use as anti-kinase therapeutics.  Oligonucleotides
CC   based on the polynucleotide sequences, i.e. probes and antisense
CC   constructs, the peptides and antibodies are claimed to be useful as
CC   tools for studying signalling cascades in cells and proteins, and for
CC   identifying inhibitors (drugs) to treat diseases and inflammatory
CC   conditions associated with abnormal kinase expression.  Diseases that
CC   are claimed to be treatable include human X-linked agammaglobulinaemia,
CC   nonspherocytic haemolytic anaemia, artherosclerosis, carcinomas,
CC   diabetes, gliomas, restenosis, cholera-based septic shock, etc.
XX
SQ   Sequence 210 BP; 74 A; 40 C; 41 G; 55 T; 0 other;

  Query Match             10.0%;  Score 210;  DB 19;  Length 210;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-32;
  Matches  210;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    844 tgaaaactatgaacctggacaaaaatcaagggccagtatcaagcacgatatatatagcta 903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 tgaaaactatgaacctggacaaaaatcaagggccagtatcaagcacgatatatatagcta 60
```

```
Qy    904 tgcagttatcacatgggaagtgttatccagaaaacagccttttgaagatgtcaccaatcc 963
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 tgcagttatcacatgggaagtgttatccagaaaacagccttttgaagatgtcaccaatcc 120

Qy    964 tttgcagataatgtatagtgtgtcacaaggacatcgacctgttattaatgaagaaagttt 1023
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 tttgcagataatgtatagtgtgtcacaaggacatcgacctgttattaatgaagaaagttt 180

Qy   1024 gccatatgatatacctcaccgagcacgtat 1053
          ||||||||||||||||||||||||||||||
Db    181 gccatatgatatacctcaccgagcacgtat 210


RESULT   8
Z61830
ID   Z61830 standard; cDNA; 1774 BP.
XX
AC   Z61830;
XX
DT   27-MAR-2000  (first entry)
XX
DE   cDNA encoding murine RIP protein kinase homologue, SEQ ID NO:403.
XX
KW   Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
KW   embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
KW   secreted; transmembrane; inflammation; cancer; neurological disease;
KW   angiogenesis; tumour vascularisation; growth disorder;
KW   developmental disorder; skin wound; hair follicle disorder;
KW   anti-inflammatory; cytostatic; neuroprotective; vulnery; ss.
XX
OS   Mus sp.
XX
PN   WO9955865-A1.
XX
PD   04-NOV-1999.
XX
PF   29-APR-1999;   99WO-NZ00051.
XX
PR   29-APR-1998;   98US-0069726.
PR   09-NOV-1998;   98US-0188930.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Strachan L,  Sleeman M,  Watson JD,  Onrust R,  Kumble A,  Murison JG;
XX
DR   WPI; 2000-072177/06.
DR   P-PSDB; Y76123.
XX
PT   Novel polynucleotides useful for the treatment of various conditions
PT   including wounds and cancer -
XX
PS   Claim 1; Page 228-229; 235pp; English.
XX
CC   The invention relates to novel nucleic acid sequences derived from rat
CC   dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
CC   and mouse embryonic skin, keratinocyte stem cells and transit amplifying
CC   cells. Polypeptides of the invention may be used to treat inflammation,
CC   cancer and neurological diseases. The proteins may be used to stimulate
CC   the growth and motility of keratinocytes, to inhibit the growth of
CC   cancer cells, to modulate angiogenesis and tumour vascularisation, to
CC   modulate skin inflammation, to modulate epithelial cell growth and to
CC   inhibit binding of HIV-1 to leukocytes. The invention may also be used to
CC   treat growth and developmental defects, skin wounds and hair follicle
CC   disorders. Sequences Z61606-Z61832 represent cDNA sequences derived from
CC   several mouse, rat or human skin cell types. Sequences Z61606-Z61649,
CC   Z61725-Z61765, Z61802-Z61811 and Z61826 encode proteins with an
CC   N-terminal signal sequence, indicating that the proteins are secreted.
CC   Sequences Z61650-Z61668, Z61766-Z61780, Z61812-Z61817 and Z61827-Z61829
CC   encode proteins with one or more putative transmembrane domains.
XX
SQ   Sequence 1774 BP; 406 A; 490 C; 526 G; 352 T; 0 other;
```

```
  Query Match             4.9%;  Score 102.4;  DB 21;  Length 1774;
  Best Local Similarity   51.8%;  Pred. No. 2.7e-11;
  Matches  372;  Conservative    0;  Mismatches  316;  Indels   30;  Gaps    5;

Qy    287 ctgcccaccattccctaccacaaactcgccgacctgcgctacctgagccgcggcgcctct 346
          |||| |||| |   |  |  | || |||| | |  |    |  || ||   |||| ||
Db     45 ctgcgcaccttcgacgccggcgaattcgcaggctgggagaaggtgggctcgggcggcttc 104

Qy    347 ggcactgtgtcgtccgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctg 406
          ||    ||||     |  ||||| |   |||||        | | ||  | |||  |  |
Db    105 gggcaggtgtacaaggtgcgccatgtgcactggaagacgtggctcgcgatcaagtgctcg 164

Qy    407 cacatccacactccgctgctcgacagtgaaagaaaggatgttttaagagaagctgaaatt 466
          | ||  |     |   |   |||||| ||| ||| |||  |  |    ||||||| | |
Db    165 cccagtc---tgcacgtcgacgacagggaacgaatggagctcctggaggaagctaagaag 221

Qy    467 ttacacaaagctagatttagttacatttttccaattttgggaatttgcaatgagcctgaa 526
           |  | |  || |  ||  | |||||| | ||  | |  || || ||| | || ||||
Db    222 atggagatggccaagttccgatacattctacctgtgtacggcatatgccaggaacctg-- 279

Qy    527 tttttgggaatagttactgaatacatgccaaatggatcattaaatgaactcctacatagg 586
              | ||  | || |  || ||||||   |  || ||  |  |  | || ||
Db    280 ----tcggcttggtcatggagtacatggagacaggctccctggagaagctgctggcctca 335

Qy    587 aaaactgaatatcctgatgttgcttggccattgagatttcgcatcctgcatgaaattgcc 646
           |  |              |  |||||    || | ||||||||| |||| || |  |||
Db    336 gagccat------------tgccttgggacctgcgctttcgcatcgtgcacgagacagcc 383

Qy    647 cttggtgtaaattacctgcacaatatgactcctcctttacttcatcatgacttgaagact 706
           | ||  | || | ||||||    ||| |||| ||  | || || |  ||| ||||| |
Db    384 gtgggcatgaacttcctgcattgcatgtctccgccactgctgcacctagacctgaagcca 443

Qy    707 cagaatatcttattggacaatgaatttcatgttaagattgcagattttggtttatcaaag 766
            ||| ||| |  ||||     | |  ||||| |||||| | || |||||  |  | |||
Db    444 gcgaacatcctgctggatgcccactaccatgtcaagatttctgactttgggctggccaag 503

Qy    767 tggcgcatgatgtccctctcacagtcacgaagtagcaaatctgcaccagaaggagggaca 826
          ||       ||||||| ||| ||    |  || |   |                || |||
Db    504 tgcaatggcatgtcccactctcatgacctcagcatggatggcctgttt------ggtaca 557

Qy    827 attatttatatgccacctgaaaactatgaacctggacaaaaatcaagggccagtatcaag 886
          ||   |||  | || || ||      |   | ||   | |       |      | |||
Db    558 atcgcttacctccctccagagcgaatt---cgtgagaagagccgcttgtttgacaccaaa 614

Qy    887 cacgatatatatagctatgcagttatcacatgggaagtgttatccagaaaacagcctttt 946
          || ||| |||| ||||  ||  || | |  ||||  ||| |  |    ||  |||| |||
Db    615 catgatgtatacagcttcgccattgtgatctggggtgtgcttacacagaagaagccattt 674

Qy    947 gaagatgtcaccaatcctttgcagataatgtatagtgtgtcacaaggacatcgacctg 1004
          | |||||  |  ||     | || || |||   |  |||  | | || || || || |
Db    675 gcagatgaaaagaacatcctacacatcatgatgaaagtggtaaagggccaccgcccag 732


RESULT   9
C99763
ID   C99763 standard; cDNA; 1774 BP.
XX
AC   C99763;
XX
DT   08-MAR-2001  (first entry)
XX
DE   Skin cell cDNA, SEQ ID NO: 403.
XX
KW   Mouse; skin cell; cytostatic; antiinflammatory; anti-HIV;
KW   nootropic; neuroprotective; vulnerary; immunomodulatory; vaccine;
KW   keratinocyte growth stimulation; cancer; angiogenesis inhibition;
KW   inflammation; neurological disease; ss.
XX
OS   Mus sp.
XX
PN   WO200069884-A2.
XX
```

```
PD   23-NOV-2000.
XX
PF   15-MAY-2000; 2000WO-NZ00075.
XX
PR   14-MAY-1999;   99US-0312283.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Watson JD,  Strachan L,  Onrust R,  Sleeman M,  Kumble KD,  Murison JG;
XX
DR   WPI; 2001-007495/01.
DR   P-PSDB; B56062.
XX
PT   New isolated polynucleotide used in the identification of genetic
PT   disorders and encoding polypeptides used for treating inflammatory
PT   disease, cancer and neurological diseases -
XX
PS   Claim 1; Page 305-306; 352pp; English.
XX
CC   The present polynucleotide encodes a polypeptide which is expressed in
CC   mammalian skin cells. The polypeptide is useful for stimulating
CC   keratinocyte growth and motility, inhibiting the growth of cancer cells,
CC   modulating angiogenesis, inhibiting angiogenesis and vascularisation of
CC   tumours, modulating skin inflammation, stimulating the growth of
CC   epithelial cells, inhibiting the binding of human immunodeficiency virus
CC   (HIV)-1 to leukocytes, and treating inflammmatory disease, cancer and
CC   neurological diseases. The polynucleotide can be used as a marker, in
CC   the identification of genetic disorders, and for the design of
CC   oligonucleotides for examining expression patterns.
XX
SQ   Sequence 1774 BP; 406 A; 490 C; 526 G; 352 T; 0 other;


  Query Match             4.9%;  Score 102.4;  DB 22;  Length 1774;
  Best Local Similarity   51.8%;  Pred. No. 2.7e-11;
  Matches  372;  Conservative    0;  Mismatches  316;  Indels   30;  Gaps    5;

Qy     287 ctgcccaccattccctaccacaaactcgccgacctgcgctacctgagccgcggcgcctct 346
           |||| |||| |   |  |  | || |||| | |  |    |  || ||   |||| ||
Db      45 ctgcgcaccttcgacgccggcgaattcgcaggctgggagaaggtgggctcgggcggcttc 104

Qy     347 ggcactgtgtcgtccgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctg 406
           ||    ||||     |  ||||| |   |||||        | | ||  | |||  |  |
Db     105 gggcaggtgtacaaggtgcgccatgtgcactggaagacgtggctcgcgatcaagtgctcg 164

Qy     407 cacatccacactccgctgctcgacagtgaaagaaaggatgttttaagagaagctgaaatt 466
           | ||  |     |   |   |||||| ||| ||| |||  |  |    |||||| | |
Db     165 cccagtc---tgcacgtcgacgacagggaacgaatggagctcctggaggaagctaagaag 221

Qy     467 ttacacaaagctagatttagttacattttttccaattttgggaatttgcaatgagcctgaa 526
            |  | |  || |  ||  | |||||| | ||  | |  || || ||| | || ||||
Db     222 atggagatggccaagttccgatacattctacctgtgtacggcatatgccaggaacctg-- 279

Qy     527 tttttgggaatagttactgaatacatgccaaatggatcattaaatgaactcctacatagg 586
               | ||  | || |  || ||||||   |  || ||  |  |  | || ||
Db     280 ----tcggcttggtcatggagtacatggagacaggctccctggagaagctgctggcctca 335

Qy     587 aaaactgaatatcctgatgttgcttggccattgagatttcgcatcctgcatgaaattgcc 646
            |  |              |  |||||    || | ||||||||| |||| || |  |||
Db     336 gagccat------------tgccttgggacctgcgctttcgcatcgtgcacgagacagcc 383

Qy     647 cttggtgtaaattacctgcacaatatgactcctcctttacttcatcatgacttgaagact 706
            | ||  | || | ||||||    ||| |||| ||  | || || |  ||| ||||| |
Db     384 gtgggcatgaacttcctgcattgcatgtctccgccactgctgcacctagacctgaagcca 443

Qy     707 cagaatatcttattggacaatgaatttcatgttaagattgcagattttggtttatcaaag 766
             ||| ||| |  ||||     | |  |||||| |||||| | || |||||  |  | |||
Db     444 gcgaacatcctgctggatgcccactaccatgtcaagatttctgactttgggctggccaag 503

Qy     767 tggcgcatgatgtccctctcacagtcacgaagtagcaaatctgcaccagaaggagggaca 826
           ||       ||||||| ||| ||    |  || |   |                || |||
Db     504 tgcaatggcatgtcccactctcatgacctcagcatggatggcctgttt------ggtaca 557

Qy     827 attatttatatgccacctgaaaactatgaacctggacaaaaatcaagggccagtatcaag 886
           ||   |||  | || || ||      |   | ||   | |       |      | |||
Db     558 atcgcttacctccctccagagcgaatt---cgtgagaagagccgcttgtttgacaccaaa 614

Qy     887 cacgatatatatagctatgcagttatcacatgggaagtgttatccagaaaacagcctttt 946
           || ||| |||| ||||  ||  || | |  ||||  ||| |  |    ||  |||| |||
Db     615 catgatgtatacagcttcgccattgtgatctggggtgtgcttacacagaagaagccattt 674

Qy     947 gaagatgtcaccaatcctttgcagataatgtatagtgtgtcacaaggacatcgacctg 1004
           | |||||  |  ||     | || || |||   |  |||  | | || || || ||  |
Db     675 gcagatgaaaagaacatcctacacatcatgatgaaagtggtaaagggccaccgcccag 732


RESULT  10
Z61161
ID   Z61161 standard; DNA; 2370 BP.
XX
AC   Z61161;
XX
DT   30-MAY-2000  (first entry)
XX
DE   DNA encoding a death associated kinase with ankyrin repeats.
XX
KW   Death associated kinase protein containing ankyrin repeats; DAKAR;
KW   kinase; quality assurance agent; shelf life; marker; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             10..2370
FT                   /*tag=  a
FT                   /transl_except= (pos:10..12, aa: Met)
XX
PN   WO200008177-A2.
XX
PD   17-FEB-2000.
XX
PF   04-AUG-1999;   99WO-US17576.
XX
PR   04-AUG-1998;   98US-0095269.
PR   11-SEP-1998;   98US-0099973.
PR   09-FEB-1999;   99US-0119353.
XX
PA   (IMMV ) IMMUNEX CORP.
XX
PI   Bird TA,  Virca GD;
XX
DR   WPI; 2000-195582/17.
DR   P-PSDB; Y69163.
XX
PT   Novel death associated kinase containing ankyrin repeats (DAKAR) used
PT   as molecular weight marker and as controls for peptide fragmentation  -
XX
PS   Claim 1; Page 9-10; 71pp; English.
XX
CC   The present sequence encodes a murine death associated kinase protein,
CC   containing ankyrin repeats, designated DAKAR. The DAKAR polynucleotides
CC   can be used to express the polypeptides, and as probes to identify
CC   nucleic acids encoding proteins having kinase activity. DAKAR
CC   polypeptides and fragmented polypeptides are used for purifying
CC   proteins, e.g. to measure protein activity; as quality assurance agents
CC   to monitor shelf life and stability of binding partner proteins; as
CC   research agents, e.g. in assays to determine protein kinase activity,
CC   to identify novel molecules involved in signal transduction pathways,
CC   and to identify therapeutic compounds which may interfere with
CC   apoptosis; as molecular weight and isoelectric focusing markers; as
CC   controls for peptide fragmentation; identification of unknown proteins,
CC   e.g. by comparison with proteins in databases; and for preparation of
CC   antibodies. The antibodies can be used in assays to detect the presence
CC   of the protein, and to purify the protein by immunoaffinity
CC   chromatography. The antibodies can also be used to block binding of
CC   the DAKAR polypeptides to their binding partners. Compounds that inhibit
CC   or enhance the kinase activity of DAKAR can be used to treat diseases
```

```
CC   characterized by overproduction or upregulated production or
CC   underproduction or downregulated production of DAKAR.
XX
SQ   Sequence 2370 BP; 536 A; 660 C; 700 G; 474 T; 0 other;

  Query Match             4.9%;  Score 102.4;  DB 21;  Length 2370;
  Best Local Similarity   51.8%;  Pred. No. 2.9e-11;
  Matches  372;  Conservative    0;  Mismatches  316;  Indels   30;  Gaps    5;

Qy     287 ctgcccaccattccctaccacaaactcgccgacctgcgctacctgagccgcggcgcctct 346
           |||| |||| |   |  |  | || |||| | |  |    |  || ||   |||| ||
Db      49 ctgcgcaccttcgacgccggcgaattcgcaggctgggagaaggtgggctcgggcggcttc 108

Qy     347 ggcactgtgtcgtccgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctg 406
           ||    ||||     |  ||||| |   ||||||        | | ||  | |||  |  |
Db     109 gggcaggtgtacaaggtgcgccatgtgcactggaagacgtggctcgcgatcaagtgctcg 168

Qy     407 cacatccacactccgctgctcgacagtgaaagaaaggatgttttaagagaagctgaaatt 466
           | ||  |     |   |   |||||| ||| ||| |||  |  |    ||||||| | |
Db     169 cccagtc---tgcacgtcgacgacagggaacgaatggagctcctggaggaagctaagaag 225

Qy     467 ttacacaaagctagatttagttacatttttccaattttgggaatttgcaatgagcctgaa 526
            |  | |  || |  ||  | |||||| | ||  | |  || || ||| | || ||||
Db     226 atggagatggccaagttccgatacattctacctgtgtacggcatatgccaggaacctg-- 283

Qy     527 tttttgggaatagttactgaatacatgccaaatggatcattaaatgaactcctacatagg 586
               | ||  | || |  || ||||||   |  || ||  |  |  | ||  ||
Db     284 ----tcggcttggtcatggagtacatggagacaggctccctggagaagctgctggcctca 339

Qy     587 aaaactgaatatcctgatgttgcttggccattgagatttcgcatcctgcatgaaattgcc 646
            |  |              |  |||||    || | |||||||| |||| || |  |||
Db     340 gagccat------------tgccttgggacctgcgctttcgcatcgtgcacgagacagcc 387

Qy     647 cttggtgtaaattacctgcacaatatgactcctcctttacttcatcatgacttgaagact 706
            | ||  | || | ||||||    ||| |||| ||  | || || |  ||| ||||| |
Db     388 gtgggcatgaacttcctgcattgcatgtctccgccactgctgcacctagacctgaagcca 447

Qy     707 cagaatatcttattggacaatgaatttcatgttaagattgcagattttggtttatcaaag 766
             ||| ||| |  ||||     | |  |||||| |||||| | || |||||  |  | |||
Db     448 gcgaacatcctgctggatgcccactaccatgtcaagatttctgactttgggctggccaag 507

Qy     767 tggcgcatgatgtccctctcacagtcacgaagtagcaaatctgcaccagaaggagggaca 826
           ||       |||||||| ||| ||    |  || |   |                || |||
Db     508 tgcaatggcatgtcccactctcatgacctcagcatggatggcctgttt------ggtaca 561

Qy     827 attatttatatgccacctgaaaactatgaacctggacaaaaatcaagggccagtatcaag 886
           ||   |||  | || || ||      |   | ||   | |       |      | |||
Db     562 atcgcttacctccctccagagcgaatt---cgtgagaagagccgcttgtttgacaccaaa 618

Qy     887 cacgatatatatagctatgcagttatcacatgggaagtgttatccagaaaacagcctttt 946
           || ||| |||| ||||  ||  || | |  ||||  ||| |  |    ||  |||| |||
Db     619 catgatgtatacagcttcgccattgtgatctggggtgtgcttacacagaagaagccattt 678

Qy     947 gaagatgtcaccaatcctttgcagataatgtatagtgtgtcacaaggacatcgacctg 1004
           | |||||  |  ||     | || || |||   |  |||   | | || || || || |
Db     679 gcagatgaaaagaacatcctacacatcatgatgaaagtggtaaagggccaccgcccag 736

RESULT  11
Z61784
ID   Z61784 standard; cDNA; 3516 BP.
XX
AC   Z61784;
XX
DT   27-MAR-2000  (first entry)
XX
DE   cDNA encoding murine protein kinase/ankyrin homologue, SEQ ID NO:257.
XX
KW   Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
KW   embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
KW   secreted; transmembrane; inflammation; cancer; neurological disease;
KW   angiogenesis; tumour vascularisation; growth disorder;
KW   developmental disorder; skin wound; hair follicle disorder;
KW   anti-inflammatory; cytostatic; neuroprotective; vulnery; ss.
XX
OS   Mus sp.
XX
PN   WO9955865-A1.
XX
PD   04-NOV-1999.
XX
PF   29-APR-1999;   99WO-NZ00051.
XX
PR   29-APR-1998;   98US-0069726.
PR   09-NOV-1998;   98US-0188930.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Strachan L,  Sleeman M,  Watson JD,  Onrust R,  Kumble A,  Murison JG;
XX
DR   WPI; 2000-072177/06.
DR   P-PSDB; Y76079.
XX
PT   Novel polynucleotides useful for the treatment of various conditions
PT   including wounds and cancer -
XX
PS   Claim 1; Page 158-159; 235pp; English.
XX
CC   The invention relates to novel nucleic acid sequences derived from rat
CC   dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
CC   and mouse embryonic skin, keratinocyte stem cells and transit amplifying
CC   cells. Polypeptides of the invention may be used to treat inflammation,
CC   cancer and neurological diseases. The proteins may be used to stimulate
CC   the growth and motility of keratinocytes, to inhibit the growth of
CC   cancer cells, to modulate angiogenesis and tumour vascularisation, to
CC   modulate skin inflammation, to modulate epithelial cell growth and to
CC   inhibit binding of HIV-1 to leukocytes. The invention may also be used to
CC   treat growth and developmental defects, skin wounds and hair follicle
CC   disorders. Sequences Z61606-Z61832 represent cDNA sequences derived from
CC   several mouse, rat or human skin cell types. Sequences Z61606-Z61649,
CC   Z61725-Z61765, Z61802-Z61811 and Z61826 encode proteins with an
CC   N-terminal signal sequence, indicating that the proteins are secreted.
CC   Sequences Z61650-Z61668, Z61766-Z61780, Z61812-Z61817 and Z61827-Z61829
CC   encode proteins with one or more putative transmembrane domains.
XX
SQ   Sequence 3516 BP; 810 A; 886 C; 988 G; 832 T; 0 other;

  Query Match             4.9%;  Score 102.4;  DB 21;  Length 3516;
  Best Local Similarity   51.8%;  Pred. No. 3.2e-11;
  Matches  372;  Conservative    0;  Mismatches  316;  Indels   30;  Gaps    5;

Qy     287 ctgcccaccattccctaccacaaactcgccgacctgcgctacctgagccgcggcgcctct 346
           |||| |||| |   |  |  | || |||| | |  |    |  || ||   |||| ||
Db      43 ctgcgcaccttcgacgccggcgaattcgcaggctgggagaaggtgggctcgggcggcttc 102

Qy     347 ggcactgtgtcgtccgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctg 406
           ||    ||||     |  ||||| |   ||||||        | | ||  | |||  |  |
Db     103 gggcaggtgtacaaggtgcgccatgtgcactggaagacgtggctcgcgatcaagtgctcg 162

Qy     407 cacatccacactccgctgctcgacagtgaaagaaaggatgttttaagagaagctgaaatt 466
           | ||  |     |   |   |||||| ||| ||| |||  |  |    ||||||| | |
Db     163 cccagtc---tgcacgtcgacgacagggaacgaatggagctcctggaggaagctaagaag 219

Qy     467 ttacacaaagctagatttagttacatttttccaattttgggaatttgcaatgagcctgaa 526
            |  | |  || |  ||  | |||||| | ||  | |  || || ||| | || ||||
Db     220 atggagatggccaagttccgatacattctacctgtgtacggcatatgccaggaacctg-- 277

Qy     527 tttttgggaatagttactgaatacatgccaaatggatcattaaatgaactcctacatagg 586
               | ||  | || |  || ||||||   |  || ||  |  |  | ||  ||
Db     278 ----tcggcttggtcatggagtacatggagacaggctccctggagaagctgctggcctca 333

Qy     587 aaaactgaatatcctgatgttgcttggccattgagatttcgcatcctgcatgaaattgcc 646
            |  |              |  |||||    || | |||||||| |||| || |  |||
Db     334 gagccat------------tgccttgggacctgcgctttcgcatcgtgcacgagacagcc 381
```

```
Qy    647 cttggtgtaaattacctgcacaatatgactcctcctttacttcatcatgacttgaagact 706
           | ||  | || | ||||||    ||| |||| ||  | || || |  ||| ||||| |
Db    382 gtgggcatgaacttcctgcattgcatgtctccgccactgctgcacctagacctgaagcca 441

Qy    707 cagaatatcttattggacaatgaatttcatgttaagattgcagattttggtttatcaaag 766
            ||| ||| |  ||||     | |  ||||| |||||| | || |||||  |  | |||
Db    442 gcgaacatcctgctggatgcccactaccatgtcaagatttctgactttgggctggccaag 501

Qy    767 tggcgcatgatgtccctctcacagtcacgaagtagcaaatctgcaccagaaggagggaca 826
          ||       ||||||| ||| ||    |  || |   |                || |||
Db    502 tgcaatggcatgtcccactctcatgacctcagcatggatggcctgttt------ggtaca 555

Qy    827 attatttatatgccacctgaaaactatgaacctggacaaaaatcaagggccagtatcaag 886
          ||   |||  | || || ||      |   | ||   | |       |      | |||
Db    556 atcgcttacctccctccagagcgaatt---cgtgagaagagccgcttgtttgacaccaaa 612

Qy    887 cacgatatatatagctatgcagttatcacatgggaagtgttatccagaaaacagcctttt 946
          || ||| |||| ||||  ||  || | |  ||||  ||| |  |    ||  |||| |||
Db    613 catgatgtatacagcttcgccattgtgatctggggtgtgcttacacagaagaagccattt 672

Qy    947 gaagatgtcaccaatcctttgcagataatgtatagtgtgtcacaaggacatcgacctg 1004
          | |||||  |  ||     | || || |||   |  |||  | | || || || || |
Db    673 gcagatgaaaagaacatcctacacatcatgatgaaagtggtaaagggccaccgcccag 730

RESULT  12
C99717
ID   C99717 standard; cDNA; 3516 BP.
XX
AC   C99717;
XX
DT   08-MAR-2001  (first entry)
XX
DE   Skin cell cDNA, SEQ ID NO: 257.
XX
KW   Mouse; skin cell; cytostatic; antiinflammatory; anti-HIV;
KW   nootropic; neuroprotective; vulnerary; immunomodulatory; vaccine;
KW   keratinocyte growth stimulation; cancer; angiogenesis inhibition;
KW   inflammation; neurological disease; ss.
XX
OS   Mus sp.
XX
PN   WO200069884-A2.
XX
PD   23-NOV-2000.
XX
PF   15-MAY-2000; 2000WO-NZ00075.
XX
PR   14-MAY-1999;   99US-0312283.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Watson JD,  Strachan L,  Onrust R,  Sleeman M,  Kumble KD,  Murison JG;
XX
DR   WPI; 2001-007495/01.
DR   P-PSDB; B55958.
XX
PT   New isolated polynucleotide used in the identification of genetic
PT   disorders and encoding polypeptides used for treating inflammatory
PT   disease, cancer and neurological diseases -
XX
PS   Claim 1; Page 216-218; 352pp; English.
XX
CC   The present polynucleotide encodes a polypeptide which is expressed in
CC   mammalian skin cells. The polypeptide is useful for stimulating
CC   keratinocyte growth and motility, inhibiting the growth of cancer cells,
CC   modulating angiogenesis, inhibiting angiogenesis and vascularisation of
CC   tumours, modulating skin inflammation, stimulating the growth of
CC   epithelial cells, inhibiting the binding of human immunodeficiency virus
CC   (HIV)-1 to leukocytes, and treating inflammmatory disease, cancer and
CC   neurological diseases. The polynucleotide can be used as a marker, in
CC   the identification of genetic disorders, and for the design of
CC   oligonucleotides for examining expression patterns.
XX
SQ   Sequence 3516 BP; 810 A; 886 C; 988 G; 832 T; 0 other;

  Query Match              4.9%;  Score 102.4;  DB 22;  Length 3516;
  Best Local Similarity   51.8%;  Pred. No. 3.2e-11;
  Matches  372;  Conservative    0;  Mismatches  316;  Indels   30;  Gaps    5;

Qy    287 ctgcccaccattccctaccacaaactcgccgacctgcgctacctgagccgcggcgcctct 346
          |||| |||| |   |  |  | || |||| | |  |    |  || ||   |||| ||
Db     43 ctgcgcaccttcgacgccggcgaattcgcaggctgggagaaggtgggctcgggcggcttc 102

Qy    347 ggcactgtgtcgtccgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctg 406
          ||    ||||     |  ||||| |   |||||        | | ||  | |||  |  |
Db    103 gggcaggtgtacaaggtgcgccatgtgcactggaagacgtggctcgcgatcaagtgctcg 162

Qy    407 cacatccacactccgctgctcgacagtgaaagaaaggatgttttaagagaagctgaaatt 466
          | ||  |     |   |   |||||| ||| ||| |||  |  |    |||||| | |
Db    163 cccagtc---tgcacgtcgacgacagggaacgaatggagctcctggaggaagctaagaag 219

Qy    467 ttacacaaagctagatttagttacatttttccaattttgggaatttgcaatgagcctgaa 526
           |  | |  || |  ||  | |||||| | ||  | |  || || ||| | || ||||
Db    220 atggagatggccaagttccgatacattctacctgtgtacggcatatgccaggaacctg-- 277

Qy    527 tttttgggaatagttactgaatacatgccaaatggatcattaaatgaactcctacatagg 586
              | ||  | || |  || ||||||   |  || ||  |  |  | || ||
Db    278 ----tcggcttggtcatggagtacatggagacaggctccctggagaagctgctggcctca 333

Qy    587 aaaactgaatatcctgatgttgcttggccattgagatttcgcatcctgcatgaaattgcc 646
           |  |              |  |||||    || | ||||||||| |||| || |  |||
Db    334 gagccat------------tgccttgggacctgcgctttcgcatcgtgcacgagacagcc 381

Qy    647 cttggtgtaaattacctgcacaatatgactcctcctttacttcatcatgacttgaagact 706
           | ||  | || | ||||||    ||| |||| ||  | || || |  ||| ||||| |
Db    382 gtgggcatgaacttcctgcattgcatgtctccgccactgctgcacctagacctgaagcca 441

Qy    707 cagaatatcttattggacaatgaatttcatgttaagattgcagattttggtttatcaaag 766
            ||| ||| |  ||||     | |  ||||| |||||| | || |||||  |  | |||
Db    442 gcgaacatcctgctggatgcccactaccatgtcaagatttctgactttgggctggccaag 501

Qy    767 tggcgcatgatgtccctctcacagtcacgaagtagcaaatctgcaccagaaggagggaca 826
          ||       ||||||| ||| ||    |  || |   |                || |||
Db    502 tgcaatggcatgtcccactctcatgacctcagcatggatggcctgttt------ggtaca 555

Qy    827 attatttatatgccacctgaaaactatgaacctggacaaaaatcaagggccagtatcaag 886
          ||   |||  | || || ||      |   | ||   | |       |      | |||
Db    556 atcgcttacctccctccagagcgaatt---cgtgagaagagccgcttgtttgacaccaaa 612

Qy    887 cacgatatatatagctatgcagttatcacatgggaagtgttatccagaaaacagcctttt 946
          || ||| |||| ||||  ||  || | |  ||||  ||| |  |    ||  |||| |||
Db    613 catgatgtatacagcttcgccattgtgatctggggtgtgcttacacagaagaagccattt 672

Qy    947 gaagatgtcaccaatcctttgcagataatgtatagtgtgtcacaaggacatcgacctg 1004
          | |||||  |  ||     | || || |||   |  |||  | | || || || || |
Db    673 gcagatgaaaagaacatcctacacatcatgatgaaagtggtaaagggccaccgcccag 730

RESULT  13
T19776
ID   T19776 standard; cDNA to mRNA; 108 BP.
XX
AC   T19776;
XX
DT   12-JUL-1996  (first entry)
XX
DE   Human gene signature HUMGS00852.
XX
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
XX
OS   Homo sapiens.
```

```
XX
PN   WO9514772-A1.
XX
PD   01-JUN-1995.
XX
PF   11-NOV-1994;   94WO-JP01916.
XX
PR   12-NOV-1993;   93JP-0355504.
XX
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
XX
PI   Matsubara K,  Okubo K;
XX
DR   WPI; 1995-206931/27.
XX
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
XX
PS   Claim 1; Page 469; 2245pp; Japanese.
XX
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
XX
SQ   Sequence 108 BP; 36 A; 14 C; 13 G; 39 T; 6 other;

  Query Match             4.9%;  Score 102;  DB 16;  Length 108;
  Best Local Similarity   97.1%;  Pred. No. 1.7e-11;
  Matches  102;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy   1842 gatcaccatctttaaatttacttcaaaataaaagcatgtaagtgactgttttcaagaag 1901
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db      1 gatcaccatctttaaatttacttcaaaataaaagcatgtaagtnactgttttcaagaag 60

Qy   1902 aaatgtgtttcataaaaggatatttatatctctgttgctttgact 1946
          ||||||||||||||||||||||||||||||||| | |||||||||
Db     61 aaatgtgtttcataaaaggatatttatatctctntngctttgact 105


RESULT  14
Z58584
ID   Z58584 standard; cDNA; 2294 BP.
XX
AC   Z58584;
XX
DT   05-JUN-2000  (first entry)
XX
DE   Mouse protein kinase Fel cDNA.
XX
KW   Fel; protein kinase; mouse; signal transduction; ss.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             3..2294
FT                   /*tag=  a
XX
```

```
PN   WO200008178-A2.
XX
PD   17-FEB-2000.
XX
PF   04-AUG-1999;   99WO-US17577.
XX
PR   04-AUG-1998;   98US-0095269.
PR   11-SEP-1998;   98US-0099973.
XX
PA   (IMMV ) IMMUNEX CORP.
XX
PI   Virca GD,  Bird TA,  Anderson DM,  Marken JS;
XX
DR   WPI; 2000-205722/18.
DR   P-PSDB; Y79154.
XX
PT   Novel murine polynucleotides encoding kinase polypeptides, used as
PT   probes to identify nucleic acids encoding proteins having kinase
PT   activity -
XX
PS   Claim 1; Fig 3; 93pp; English.
XX
CC   The present sequence is that of Genesis clone 971025TRAM004820HT
CC   (Fel) DNA.  The translation of the clone (frame 3) is provided in
CC   Y79154.  The polypeptide has kinase activity.  The invention
CC   relates to purified murine polypeptides (see Y79152-56) that
CC   have kinase function and isolated nucleic acids encoding them (see
CC   Z58582-86).  The nucleic acids can be used to express the
CC   polypeptides, as probes to identify nucleic acids encoding proteins
CC   having kinase activity, and in assays to identify chromosomes, map
CC   murine genes, and study tumours.  Vectors comprising the novel
CC   nucleic acids and host cells (bacterial, yeast, plant, insect or
CC   animal) transfected or transduced with the vectors are claimed.
CC   The kinase polypeptides and their fragments are used as mol.wt. and
CC   isoelectric focusing markers, and as controls for peptide
CC   fragmentation.  They also have a number of therapeutic uses, as
CC   kinases play a central role in cellular signal transduction.
XX
SQ   Sequence 2294 BP; 526 A; 639 C; 666 G; 463 T; 0 other;

  Query Match             4.4%;  Score 93;  DB 21;  Length 2294;
  Best Local Similarity   51.9%;  Pred. No. 2e-09;
  Matches  350;  Conservative    0;  Mismatches  295;  Indels   30;  Gaps    5;

Qy    330 tgagccgcggcgcctctggcactgtgtcgtccgcccgccacgcagactggcgcgtccagg 389
          || ||   |||| ||  ||    ||||     |  ||||| |   |||||        |
Db     16 tgggctcgggcggcttcgggcaggtgtacaaggtgcgccatgtgcactggaagacgtggc 75

Qy    390 tggccgtgaagcacctgcacatccacactccgctgctcgacagtgaaagaaaggatgttt 449
          | ||  | |||  |  || ||  |     |   |   |||||| ||| ||| |||  |
Db     76 tcgcgatcaagtgctcgcccagtc---tgcacgtcgacgacagggaacgaatggagctcc 132

Qy    450 taagagaagctgaaattttacacaaagctagatttagttacatttttccaattttgggaa 509
          |    |||||| | |   |  | |  || |  ||  | |||||| | ||  | |  || |
Db    133 tggaggaagctaagaagatggagatggccaagttccgatacattctacctgtgtacggca 192

Qy    510 tttgcaatgagcctgaatttttgggaatagttactgaatacatgccaaatggatcattaa 569
          | ||| | || ||||      | ||  | || |  || ||||||   |  || ||  |
Db    193 tatgccaggaacctg------tcggcttggtcatggagtacatggagacaggctccctgg 246

Qy    570 atgaactcctacataggaaaactgaatatcctgatgttgcttggccattgagatttcgca 629
          |  | || ||        |  |              |  |||||    || | |||||||
Db    247 agaagctgctggcctcagagcc------------attgccttgggacctgcgctttcgca 294

Qy    630 tcctgcatgaaattgcccttggtgtaaattacctgcacaatatgactcctcctttacttc 689
          || |||| || |  ||| | ||  | || | ||||||    ||| |||| ||  | || |
Db    295 tcgtgcacgagacagccgtgggcatgaacttcctgcattgcatgtctccgccactgctgc 354

Qy    690 atcatgacttgaagactcagaatatcttattggacaatgaatttcatgttaagattgcag 749
          | |  ||| ||||| |   ||| ||| |  ||||     | |  ||||| |||||| | |
Db    355 acctagacctgaagccagcgaacatcctgctggatgcccactaccatgtcaagatttctg 414
```

```
Qy     750 attttggtttatcaaagtggcgcatgatgtccctctcacagtcacgaagtagcaaatctg 809
           | |||||  |  | |||||       ||||||| ||| ||    |  || |   |
Db     415 actttgggctggccaagtgcaatggcatgtcccactctcatgacctcagcatggatggcc 474

Qy     810 caccagaaggagggacaattatttatatgccacctgaaaactatgaacctggacaaaaat 869
                      || |||||   |||  | || || ||      |   | ||   | |
Db     475 tgttt------ggtacaatcgcttacctccctccagagcgaatt---cgtgagaagagcc 525

Qy     870 caagggccagtatcaagcacgatatatatagctatgcagttatcacatgggaagtgttat 929
               |      | ||| || ||| |||| ||||  ||  || | |  ||||  ||| |
Db     526 gcttgtttgacaccaaacatgatgtatacagcttcgccattgtgatctggggtgtgctta 585

Qy     930 ccagaaaacagccttttgaagatgtcaccaatcctttgcagataatgtatagtgtgtcac 989
           |    ||  |||| |||| |||||  |  ||     | || || |||   |  |||  |
Db     586 cacagaagaagccatttgcagatgaaaagaacatcctacacatcatgatgaaagtggtaa 645

Qy     990 aaggacatcgacctg 1004
           | || || || || |
Db     646 agggccaccgcccag 660

RESULT  15
Q98816
ID   Q98816 standard; cDNA; 3033 BP.
XX
AC   Q98816;
XX
DT   19-JAN-1996  (first entry)
XX
DE   Arabidopsis CTR1 cDNA.
XX
KW   CTR1; constitutive triple response; transgenic plant; ethylene;
KW   stress tolerance; crop improvement; ss.
XX
OS   Arabidopsis thaliana ecotype Columbia.
XX
FH   Key             Location/Qualifiers
FT   CDS             118..2583
FT                   /*tag= a
XX
PN   US5444166-A.
XX
PD   22-AUG-1995.
XX
PF   10-AUG-1992;   92US-0928464.
XX
PR   12-JAN-1993;   93US-0003311.
PR   10-AUG-1992;   92US-0928464.
XX
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Ecker JR,  Kieber JJ;
XX
DR   WPI; 1995-302162/39.
DR   P-PSDB; R80574.
XX
PT   New constitutive triple response genes and mutants - isolated from
PT   Arabidopsis thaliana, used to produce transgenic plants with
PT   improved properties
XX
PS   Claim 1; Column 21-28; 41pp; English.
XX
CC   The CTR1 gene was mapped to an interval between 2 RFLPs in
CC   chromosome 5 of A. thaliana and a chromosome walk in this area was
CC   initiated using a YAC library.  2 Clones were isolated and used to
CC   probe a cDNA library derived from polyA RNA from 3-day-old dark-
CC   grown, ethylene-treated A. thaliana seedlings.  The cDNA for the
CC   CTR1 gene is given in Q98816.  Expression of the CTR1 gene in
CC   transgenic plants results in a dominant ethylene-insensitive
CC   phenotype.
XX
SQ   Sequence 3033 BP; 797 A; 605 C; 751 G; 880 T; 0 other;
```

```
  Query Match             3.9%;  Score 81.4;  DB 16;  Length 3033;
  Best Local Similarity   50.1%;  Pred. No. 4.1e-07;
  Matches  202;  Conservative    0;  Mismatches  201;  Indels    0;  Gaps    0;

Qy     360 ccgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctgcacatccacactc 419
           | | || ||  || || ||||  | |  ||  |  |    | |  | | |||  |
Db    1808 ctgtccaccgtgctgagtggcatggctcggatgttgctgtgaaaattctcatggagcaag 1867

Qy     420 cgctgctcgacagtgaaagaaaggatgttttaagagaagctgaaattttacacaaagcta 479
              | |  |           || ||  | |||||||| | ||  ||  |  |      |
Db    1868 acttccatgctgagcgtgttaatgagttcttaagagaggttgcgataatgaaacgccttc 1927

Qy     480 gatttagttacatttttccaattttgggaatttgcaatgagcctgaatttttgggaatag 539
           |      | ||||| |||   |  ||||      || | | |||  |  ||||  |||||
Db    1928 gccaccctaacattgttctcttcatgggtgcggtcactcaacctccaaatttgtcaatag 1987

Qy     540 ttactgaatacatgccaaatggatcattaaatgaactcctacataggaaaactgaatatc 599
           | || |||||  || |||  ||    ||| |   |||  | ||||  |     | |
Db    1988 tgacagaatatttgtcaagaggtagtttatacagacttttgcataaaagtggagcaaggg 2047

Qy     600 ctgatgttgcttggccattgagatttcgcatcctgcatgaaattgcccttggtgtaaatt 659
              |  | | |  |  |    |  |  | ||     ||||  | ||    ||  | ||||
Db    2048 agcaattagatgagagacgtcgcctgagtatggcttatgatgtggctaagggaatgaatt 2107

Qy     660 acctgcacaatatgactcctcctttacttcatcatgacttgaagactcagaatatcttat 719
           | || ||||||   | ||||||  |  | |||   ||  | ||  |||  ||  | ||
Db    2108 atcttcacaatcgcaatcctccaattgtgcatagagatctaaaatctccaaacttattgg 2167

Qy     720 tggacaatgaatttcatgttaagattgcagattttggtttatc 762
           | |||||  ||| |   || ||| ||   ||||||||| | ||
Db    2168 ttgacaaaaaatatacagtcaaggtttgtgattttggtctctc 2210


Search completed: June 13, 2001, 17:37:08
Job time: 5922 sec
```

```
                    GenCore version 4.5
             Copyright (c) 1993 - 2000  Compugen Ltd.


OM nucleic - nucleic search, using sw model

Run on:          June 13, 2001, 15:03:41 ; Search time 3516.74 Seconds
                                             (without alignments)
                                             8798.208 Million cell updates/sec

Title:           US-09-445-223-2
Perfect score:   2098
Sequence:        1 ggccattatggatggatggg..........aaaaaaaaaaaaaaaaaaa 2098

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        1283235 seqs, 7373929652 residues

Total number of hits satisfying chosen parameters:      2566470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       GenEmbl:*
               1:  gb_ba1:*
               2:  gb_ba2:*
               3:  gb_ba3:*
               4:  gb_in1:*
               5:  gb_in2:*
               6:  gb_in3:*
               7:  gb_om:*
               8:  gb_ov:*
               9:  gb_pat1:*
               10:  gb_pat2:*
               11:  gb_ph:*
               12:  gb_pl1:*
               13:  gb_pl2:*
               14:  gb_pl3:*
               15:  gb_pl4:*
               16:  em_ba1:*
               17:  em_ba2:*
               18:  em_fun:*
               19:  em_htgo_hum:*
               20:  em_htgo_inv:*
               21:  em_htgo_rod:*
               22:  em_htg_hum1:*
               23:  em_htg_hum2:*
               24:  em_htg_hum3:*
               25:  em_htg_hum4:*
               26:  em_htg_hum5:*
               27:  em_htg_hum6:*
               28:  em_htg_hum7:*
               29:  em_htg_hum8:*
               30:  em_htg_inv1:*
               31:  em_htg_inv2:*
               32:  em_htg_other:*
               33:  em_htg_rod:*
               34:  em_hum1:*
               35:  em_hum2:*
               36:  em_hum3:*
               37:  em_hum4:*
               38:  em_hum5:*
               39:  em_hum6:*
               40:  em_hum7:*
               41:  em_in:*
               42:  em_om:*
               43:  em_or:*
```

```
               44:  em_ov:*
               45:  em_pat:*
               46:  em_ph:*
               47:  em_pl:*
               48:  em_ro:*
               49:  em_sts:*
               50:  em_sy:*
               51:  em_un:*
               52:  em_vi:*
               53:  gb_sts1:*
               54:  gb_sts2:*
               55:  gb_sts3:*
               56:  gb_sy:*
               57:  gb_un:*
               58:  gb_vi1:*
               59:  gb_vi2:*
               60:  gb_htg1:*
               61:  gb_htg2:*
               62:  gb_htg3:*
               63:  gb_htg4:*
               64:  gb_htg5:*
               65:  gb_htg6:*
               66:  gb_htg7:*
               67:  gb_htg8:*
               68:  gb_htg9:*
               69:  gb_htg10:*
               70:  gb_htg11:*
               71:  gb_htg12:*
               72:  gb_htg13:*
               73:  gb_htg14:*
               74:  gb_htg15:*
               75:  gb_htg16:*
               76:  gb_htg17:*
               77:  gb_htg18:*
               78:  gb_htg19:*
               79:  gb_htg20:*
               80:  gb_htg21:*
               81:  gb_htg22:*
               82:  gb_htg23:*
               83:  gb_htg24:*
               84:  gb_htg25:*
               85:  gb_pr1:*
               86:  gb_pr2:*
               87:  gb_pr3:*
               88:  gb_pr4:*
               89:  gb_pr5:*
               90:  gb_pr6:*
               91:  gb_pr7:*
               92:  gb_pr8:*
               93:  gb_pr9:*
               94:  gb_ro1:*
               95:  gb_ro2:*

   Pred. No. is the number of results predicted by chance to have a
   score greater than or equal to the score of the result being printed,
   and is derived by analysis of the total score distribution.

                                 SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                     Description
-------------------------------------------------------------------------------
    1  2094.4   99.8    2098  9  A82777                 A82777 Sequence 2
    2  2017.8   96.2    2501 88  AF027706                AF027706 Homo sapi
    3  1815.2   86.5    1902 88  AF064824                AF064824 Homo sapi
    4  1616.6   77.1    1623 88  AF078530                AF078530 Homo sapi
    5   516.6   24.6  116650 85  AC004003                AC004003 Homo sapi
    6   516.6   24.6  320250 88  AF117829                AF117829 Homo sapi
    7     210   10.0     210  9  AR044139               AR044139 Sequence
    8   102.4    4.9    3559 94  AF302127                AF302127 Mus muscu
    9    97.6    4.7    3879 92  HSA278016               AJ278016 Homo sapi
   10    81.4    3.9    3033 10  I14046                  I14046 Sequence 1
   11    81.4    3.9    3033 10  I35764                  I35764 Sequence 1
```

```
    12     81.4     3.9    3033   13  ATHCTR1A               L08789 Arabidopsis
    13     78.2     3.7    2767   12  AF110519               AF110519 Lycopersi
    14     78.2     3.7    2829   12  AF096250               AF096250 Lycopersi
    15     78.2     3.7    2836   12  AF110518               AF110518 Lycopersi
    16     78.2     3.7    2917   14  LETYRKINA              Y13273 Lycopersico
c   17     68.2     3.3  310779   60  AC005140               AC005140 Plasmodiu
c   18     68.2     3.3  321003   84  PFMAL4P3               AL035476 Plasmodiu
c   19     66.6     3.2  256172   60  AC005139               AC005139 Plasmodiu
    20     66.2     3.2     527   13  AF230506               AF230506 Oryza sat
    21     65.8     3.1     759   53  CNS06QXV               AL411257 T7 end of
c   22     65.6     3.1  192929   60  AC005505               AC005505 Plasmodiu
    23     65.2     3.1   80442   15  T24D18                 AC010924 Arabidops
c   24     64.6     3.1   80402   12  AB025636               AB025636 Arabidops
c   25     64.6     3.1  113567   12  AF058826               AF058826 Arabidops
    26     64.4     3.1   88037   84  PFMAL13P8              AL096782 Plasmodiu
c   27       64     3.1  131473   12  AC005171               AC005171 Arabidops
c   28     63.6     3.0  163443   60  AC006280               AC006280 Plasmodiu
    29     63.2     3.0    1833   94  AF178953               AF178953 Mus muscu
    30     62.6     3.0  156060   60  AC004153               AC004153 Plasmodiu
    31     62.4     3.0  321003   84  PFMAL4P3               AL035476 Plasmodiu
    32     61.8     2.9    1554   9  AR084102               AR084102 Sequence
    33     61.8     2.9    6372   6  DDICNPA                M23449 Dictyoseliu
c   34     61.4     2.9    5762   6  PFAHSP86A              L34027 Plasmodium
c   35     61.4     2.9   47573   5  AF030694               AF030694 Plasmodiu
c   36     61.2     2.9  214094   78  AC087188               AC087188 Homo sapi
    37       61     2.9     895   53  CNS06FSV               AL396821 T7 end of
    38     60.2     2.9    1439   5  AF038050               AF038050 Trichoplu
    39     60.2     2.9    2358   7  AF254119               AF254119 Ovis arie
    40       60     2.9    2508   9  AX056437               AX056437 Sequence
    41       60     2.9    3001   88  AF116826               AF116826 Homo sapi
c   42       60     2.9   90149   12  AC010793               AC010793 Genomic s
    43       60     2.9  175406   78  AC087128               AC087128 Mus muscu
c   44     59.8     2.9  196490   60  AC005507               AC005507 Plasmodiu
    45     59.6     2.8    2982   6  DDISGSPA               M33862 Dictyosteli


                              ALIGNMENTS


RESULT   1
A82777
LOCUS       A82777       2098 bp    DNA             PAT       21-JAN-2000
DEFINITION  Sequence 2 from Patent WO9855507.
ACCESSION   A82777
VERSION     A82777.1  GI:6732464
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2098)
  AUTHORS   Boldin,M. and Wallach,D.
  TITLE     MODULATORS OF INTRACELLULAR INFLAMMATION, CELL DEATH AND CELL
            SURVIVAL PATHWAYS
  JOURNAL   Patent: WO 9855507-A 2 10-DEC-1998;
            BOLDIN MARK (IL); WALLACH DAVID (IL)
FEATURES             Location/Qualifiers
     source          1..2098
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      649 a    452 c    449 g    539 t       9 others
ORIGIN


  Query Match              99.8%;  Score 2094.4;  DB 9;  Length 2098;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2098;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ggccattatggatggatgggcggcgctacggcgttggcaccagtctctagaaaagaagtc 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCCATTATGGATGGATGGGCGGCGCTACGGCGTTGGCACCAGTCTCTAGAAAAGAAGTC 60

Qy         61 agctctggttcggagaagcagcggctggcgtgggccatccggggaatgggcgccctcgtg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         61 AGCTCTGGTTCGGAGAAGCAGCGGCTGGCGTGGGCCATCCGGGGAATGGGCGCCCTCGTG 120

Qy        121 acctagtgttgcggggcaaaaagggtcttgccggcctcgctcgtgcaggggcgtatctgg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCTAGTGTTGCGGGGCAAAAAGGGTCTTGCCGGCCTCGCTCGTGCAGGGGCGTATCTGG 180

Qy        181 gcgcctgagcgcggcgtgggagccttgggagccgccgcagcaggggcacacccggaacc 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCGCCTGAGCGCGGCGTGGGAGCCTTGGGAGCCGCCGCAGCAGGGGCACACCCGGAACC 240

Qy        241 ggcctgagcgcccgggaccatgaacggggaggccatctgcagcgccctgcccaccattcc 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGCCTGAGCGCCCGGGACCATGAACGGGGAGGCCATCTGCAGCGCCCTGCCCACCATTCC 300

Qy        301 ctaccacaaactcgccgacctgcgctacctgagccgcggcgcctctggcactgtgtcgtc 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTACCACAAACTCGCCGACCTGCGCTACCTGAGCCGCGGCGCCTCTGGCACTGTGTCGTC 360

Qy        361 cgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctgcacatccacactcc 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGCCCGCCACGCAGACTGGCGCGTCCAGGTGGCCGTGAAGCACCTGCACATCCACACTCC 420

Qy        421 gctgctcgacagtgaaagaaaggatgttttaagagaagctgaaattttacacaaagctag 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTGCTCGACAGTGAAAGAAAGGATGTTTTAAGAGAAGCTGAAATTTTACACAAAGCTAG 480

Qy        481 atttagttacatttttccaattttgggaatttgcaatgagcctgaatttttgggaatagt 540.
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATTTAGTTACATTTTTCCAATTTTGGGAATTTGCAATGAGCCTGAATTTTTGGGAATAGT 540

Qy        541 tactgaatacatgccaaatggatcattaaatgaactcctacataggaaaactgaatatcc 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TACTGAATACATGCCAAATGGATCATTAAATGAACTCCTACATAGGAAAACTGAATATCC 600

Qy        601 tgatgttgcttggccattgagatttcgcatcctgcatgaaattgcccttggtgtaaatta 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGATGTTGCTTGGCCATTGAGATTTCGCATCCTGCATGAAATTGCCCTTGGTGTAAATTA 660

Qy        661 cctgcacaatatgactcctcctttacttcatcatgacttgaagactcagaatatcttatt 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGCACAATATGACTCCTCCTTTACTTCATCATGACTTGAAGACTCAGAATATCTTATT 720

Qy        721 ggacaatgaatttcatgttaagattgcagattttggtttatcaaagtggcgcatgatgtc 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGACAATGAATTTCATGTTAAGATTGCAGATTTTGGTTTATCAAAGTGGCGCATGATGTC 780

Qy        781 cctctcacagtcacgaagtagcaaatctgcaccagaaggagggacaattatttatatgcc 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CCTCTCACAGTCACGAAGTAGCAAATCTGCACCAGAAGGAGGGACAATTATTTATATGCC 840

Qy        841 acctgaaaactatgaacctggacaaaaatcaagggccagtatcaagcacgatatatatag 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACCTGAAAACTATGAACCTGGACAAAAATCAAGGGCCAGTATCAAGCACGATATATATAG 900

Qy        901 ctatgcagttatcacatgggaagtgttatccagaaaacagcctttgaagatgtcaccaa 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTATGCAGTTATCACATGGGAAGTGTTATCCAGAAAACAGCCTTTTGAAGATGTCACCAA 960

Qy        961 tcctttgcagataatgtatagtgtgtcacaaggacatcgacctgttattaatgaagaaag 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCCTTTGCAGATAATGTATAGTGTGTCACAAGGACATCGACCTGTTATTAATGAAGAAAG 1020

Qy       1021 tttgccatatgatatacctcaccgagcacgtatgatctctctaatagaaagtggatgggc 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTTGCCATATGATATACCTCACCGAGCACGTATGATCTCTCTAATAGAAAGTGGATGGGC 1080

Qy       1081 acaaaatccagatgaaagaccatctttcttaaaatgtttaatagaacttgaaccagtttt 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACAAAATCCAGATGAAAGACCATCTTTCTTAAAATGTTTAATAGAACTTGAACCAGTTTT 1140

Qy       1141 gagaacatttgaagagataacttttcttgaagctgttattcagctaaagaaaacaaagtt 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GAGAACATTTGAAGAGATAACTTTTCTTGAAGCTGTTATTCAGCTAAAGAAAACAAAGTT 1200
```

```
Qy   1201 acagagtgtttcaagtgccattcacctatgtgacaagaagaaaatggaattatctctgaa 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ACAGAGTGTTTCAAGTGCCATTCACCTATGTGACAAGAAGAAAATGGAATTATCTCTGAA 1260

Qy   1261 catacctgtaaatcatggtccacaagaggaatcatgtggatcctctcagctccatgaaaa 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CATACCTGTAAATCATGGTCCACAAGAGGAATCATGTGGATCCTCTCAGCTCCATGAAAA 1320

Qy   1321 tagtggttctcctgaaacttcaaggtccctgccagctcctcaagacaatgatttttatc 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TAGTGGTTCTCCTGAAACTTCAAGGTCCCTGCCAGCTCCTCAAGACAATGATTTTTATC 1380

Qy   1381 tagaaaagctcaagactgttattttatgaagctgcatcactgtcctggaaatcacagttg 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TAGAAAAGCTCAAGACTGTTATTTTATGAAGCTGCATCACTGTCCTGGAAATCACAGTTG 1440

Qy   1441 ggatagcaccatttctggatctcaaagggctgcattctgtgatcacaagaccactccatg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GGATAGCACCATTTCTGGATCTCAAAGGGCTGCATTCTGTGATCACAAGACCACTCCATG 1500

Qy   1501 ctcttcagcaataataaatccactctcaactgcaggaaactcagaacgtctgcagcctgg 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CTCTTCAGCAATAATAAATCCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAGCCTGG 1560

Qy   1561 tatagcccagcagtggatccagagcaaaagggaagacattgtgaaccaaatgacagaagc 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 TATAGCCCAGCAGTGGATCCAGAGCAAAAGGGAAGACATTGTGAACCAAATGACAGAAGC 1620

Qy   1621 ctgccttaaccagtcgctagatgcccttctgtccagggacttgatcatgaaagaggacta 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 CTGCCTTAACCAGTCGCTAGATGCCCTTCTGTCCAGGGACTTGATCATGAAAGAGGACTA 1680

Qy   1681 tgaacttgttagtaccaagcctacaaggacctcaaaagtcagacaattactagacactac 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TGAACTTGTTAGTACCAAGCCTACAAGGACCTCAAAAGTCAGACAATTACTAGACACTAC 1740

Qy   1741 tgacatccaaggagaagaatttgccaaagttatagtacaaaaattgaaagataacaaaca 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 TGACATCCAAGGAGAAGAATTTGCCAAAGTTATAGTACAAAAATTGAAAGATAACAAACA 1800

Qy   1801 aatgggtcttcagccttacccggaaatacttgtggtttctagatcaccatctttaaattt 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 AATGGGTCTTCAGCCTTACCCGGAAATACTTGTGGTTTCTAGATCACCATCTTTAAATTT 1860

Qy   1861 acttcaaaataaaagcatgtaagtgactgtttttcaagaagaaatgtgtttcataaaagg 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 ACTTCAAAATAAAAGCATGTAAGTGACTGTTTTTCAAGAAGAAATGTGTTTCATAAAAGG 1920

Qy   1921 atatttatatctctgttgctttgactttttttatataaaatccgtgagtattaaagcttw 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 ATATTTATATCTCTGTTGCTTTGACTTTTTTTATATAAAATCCGTGAGTATTAAAGCTTW 1980

Qy   1981 awwraargktctttsrktaaatattagtctccctccatgacactgcagtatttttttaa 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 AWWRAARGKTCTTTSRKTAAATATTAGTCTCCCTCCATGACACTGCAGTATTTTTTTAA 2040

Qy   2041 ttaatacaagtaaaaagttgaatttgaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2098
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TTAATACAAGTAAAAAGTTGAATTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2098

RESULT   2
AF027706
LOCUS       AF027706     2501 bp    mRNA            PRI       10-SEP-1998
DEFINITION  Homo sapiens serine/threonine kinase RICK (RICK) mRNA, complete
            cds.
ACCESSION   AF027706
VERSION     AF027706.1  GI:3123886
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2501)
  AUTHORS   Inohara,N., del Peso,L., Koseki,T., Chen,S. and Nunez,G.
  TITLE     RICK, a novel protein kinase containing a caspase recruitment
            domain, interacts with CLARP and regulates CD95-mediated apoptosis
  JOURNAL   J. Biol. Chem. 273 (20), 12296-12300 (1998)
  MEDLINE   98241596
REFERENCE   2  (bases 1 to 2501)
  AUTHORS   Inohara,N., Koseki,T., Chen,S., del Peso,L. and Nunez,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-OCT-1997) Dept. Pathlogy, Comprehensive Cancer &
            Geriatrics Center, 4-131 CCGC 1500 E. Medical Center Dr, Ann Arbor,
            MI 48109, USA
FEATURES             Location/Qualifiers
     source          1. .2501
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8q21"
     gene            1. .2501
                     /gene="RICK"
     CDS             225. .1847
                     /gene="RICK"
                     /function="regulates apoptosis"
                     /note="CARD domain protein; phosphorylated protein"
                     /codon_start=1
                     /product="serine/threonine kinase RICK"
                     /protein_id="AAC34970.1"
                     /db_xref="GI:3123887"
                     /translation="MNGEAICSALPTIPYHKLADLRYLSRGASGTVSSARHADWRVQV
                     AVKHLHIHTPLLDSERKDVLREAEILHKARFSYILPILGICNEPEFLGIVTEYMPNGS
                     LNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHV
                     KIADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVI
                     TWEVLSRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIPHRARMISLIESGWAQN
                     PDERPSFLKCLIELEPVLRTFEEITFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLN
                     IPVNHGPQEESCGSSQLHENSGSPETSRSLPAPQDNDFLSRKAQDCYFMKLHHCPGNH
                     SWDSTISGSQRAAFCDHKTTPCSSAIINPLSTAGNSERLQPGIAQQWIQSKREDIVNQ
                     MTEACLNQSLDALLSRDLIMKEDYELVSTKPTRTSKVRQLLDTTDIQGEEFAKVIVQK
                     LKDNKQMGLQPYPEILVVSRSPSLNLLQNKSM"
     polyA_signal    1833. .1838
                     /gene="RICK"
     polyA_site      1869
                     /gene="RICK"
     polyA_signal    2479. .2484
                     /gene="RICK"
     polyA_site      2501
                     /gene="RICK"
BASE COUNT      768 a    535 c    499 g    699 t
ORIGIN


  Query Match             96.2%;  Score 2017.8;  DB 88;  Length 2501;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches 2016;  Conservative    9;  Mismatches    6;  Indels    0;  Gaps    0;

Qy     36 ggcaccagtctctagaaaagaagtcagctctggttcggagaagcagcggctggcgtgggc 95
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGCACCAGTCTCTAGAAAAGAAGTCAGCTCTGGTTCGGAGAAGCAGCGGCTGGCGTGGGC 60

Qy     96 catccggggaatgggcgccctcgtgacctagtgttgcggggcaaaaagggtcttgccggc 155
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CATCCGGGGAATGGGCGCCCTCGTGACCTAGTGTTGCGGGGCAAAAAGGGTCTTGCCGGC 120

Qy    156 ctcgctcgtgcaggggcgtatctgggcgcctgagcgcggcgtgggagccttgggagccgc 215
          ||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db    121 CTCGCTCGTGCAGGGGCGTATCTGGGCGCCTGAGCGCGCAGTGGGAGCCTTGGGAGCCGC 180

Qy    216 cgcagcaggggcacacccggaaccggcctgagcgcccgggaccatgaacggggaggcca 275
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CGCAGCAGGGGCACACCCGGAACCGGCCTGAGCGCCCGGGACCATGAACGGGGAGGCCA 240

Qy    276 tctgcagcgccctgcccaccattccctaccacaaactcgccgacctgcgctacctgagcc 335
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    241 TCTGCAGCGCCCTGCCCACCATTCCCTACCACAAACTCGCCGACCTGCGCTACCTGAGCC 300

Qy    336 gcggcgcctctggcactgtgtcgtccgcccgccacgcagactggcgcgtccaggtggccg 395
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GCGGCGCCTCTGGCACTGTGTCGTCCGCCCGCCACGCAGACTGGCGCGTCCAGGTGGCCG 360

Qy    396 tgaagcacctgcacatccacactccgctgctcgacagtgaaagaaaggatgttttaagag 455
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    361 TGAAGCACCTGCACATCCACACTCCGCTGCTCGACAGTGAAAGAAAGGATGTCTTAAGAG 420

Qy    456 aagctgaaattttacacaaagctagatttagttacattttccaattttgggaatttgca 515
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    421 AAGCTGAAATTTTACACAAAGCTAGATTTAGTTACATTCTTCCAATTTTGGGAATTTGCA 480

Qy    516 atgagcctgaatttttgggaatagttactgaatacatgccaaatggatcattaaatgaac 575
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ATGAGCCTGAATTTTTGGGAATAGTTACTGAATACATGCCAAATGGATCATTAAATGAAC 540

Qy    576 tcctacataggaaaactgaatatcctgatgttgcttggccattgagatttcgcatcctgc 635
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TCCTACATAGGAAAACTGAATATCCTGATGTTGCTTGGCCATTGAGATTTCGCATCCTGC 600

Qy    636 atgaaattgcccttggtgtaaattacctgcacaatatgactcctcctttacttcatcatg 695
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ATGAAATTGCCCTTGGTGTAAATTACCTGCACAATATGACTCCTCCTTTACTTCATCATG 660

Qy    696 acttgaagactcagaatatcttattggacaatgaatttcatgttaagattgcagattttg 755
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ACTTGAAGACTCAGAATATCTTATTGGACAATGAATTTCATGTTAAGATTGCAGATTTTG 720

Qy    756 gtttatcaaagtggcgcatgatgtccctctcacagtcacgaagtagcaaatctgcaccag 815
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GTTTATCAAAGTGGCGCATGATGTCCCTCTCACAGTCACGAAGTAGCAAATCTGCACCAG 780

Qy    816 aaggagggacaattatttatatgccacctgaaaactatgaacctggacaaaaatcaaggg 875
          |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db    781 AAGGAGGGACAATTATCTATATGCCACCTGAAAACTATGAACCTGGACAAAAATCAAGGG 840

Qy    876 ccagtatcaagcacgatatatatagctatgcagttatcacatgggaagtgttatccagaa 935
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CCAGTATCAAGCACGATATATATAGCTATGCAGTTATCACATGGGAAGTGTTATCCAGAA 900

Qy    936 aacagccttttgaagatgtcaccaatcctttgcagataatgtatagtgtgtcacaaggac 995
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 AACAGCCTTTTGAAGATGTCACCAATCCTTTGCAGATAATGTATAGTGTGTCACAAGGAC 960

Qy    996 atcgacctgttattaatgaagaaagtttgccatatgatatacctcaccgagcacgtatga 1055
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ATCGACCTGTTATTAATGAAGAAAGTTTGCCATATGATATACCTCACCGAGCACGTATGA 1020

Qy   1056 tctctctaatagaaagtggatgggcacaaaatccagatgaaagaccatctttcttaaaat 1115
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TCTCTCTAATAGAAAGTGGATGGGCACAAAATCCAGATGAAAGACCATCTTTCTTAAAAT 1080

Qy   1116 gtttaatagaacttgaaccagttttgagaacatttgaagagataactttttcttgaagctg 1175
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GTTTAATAGAACTTGAACCAGTTTTGAGAACATTTGAAGAGATAACTTTTCTTGAAGCTG 1140

Qy   1176 ttattcagctaaagaaaacaaagttacagagtgtttcaagtgccattcacctatgtgaca 1235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTATTCAGCTAAAGAAAACAAAGTTACAGAGTGTTTCAAGTGCCATTCACCTATGTGACA 1200

Qy   1236 agaagaaaatggaattatctctgaacatacctgtaaatcatggtccacaagaggaatcat 1295
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AGAAGAAAATGGAATTATCTCTGAACATACCTGTAAATCATGGTCCACAAGAGGAATCAT 1260

Qy   1296 gtggatcctctcagctccatgaaaatagtggttctcctgaaacttcaaggtccctgccag 1355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTGGATCCTCTCAGCTCCATGAAAATAGTGGTTCTCCTGAAACTTCAAGGTCCCTGCCAG 1320

Qy   1356 ctcctcaagacaatgattttttatctagaaaagctcaagactgttattttatgaagctgc 1415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CTCCTCAAGACAATGATTTTTTATCTAGAAAAGCTCAAGACTGTTATTTTATGAAGCTGC 1380

Qy   1416 atcactgtcctggaaatcacagttgggatagcaccatttctggatctcaaagggctgcat 1475
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   1381 ATCACTGTCCTGGAAATCACAGTTGGGATAGCACCATTTCTGGTTCTCAAAGGGCTGCAT 1440

Qy   1476 tctgtgatcacaagaccactccatgctcttcagcaataataaatccactctcaactgcag 1535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TCTGTGATCACAAGACCACTCCATGCTCTTCAGCAATAATAAATCCACTCTCAACTGCAG 1500

Qy   1536 gaaactcagaacgtctgcagcctggtatagcccagcagtggatccagagcaaaagggaag 1595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GAAACTCAGAACGTCTGCAGCCTGGTATAGCCCAGCAGTGGATCCAGAGCAAAAGGGAAG 1560

Qy   1596 acattgtgaaccaaatgacagaagcctgccttaaccagtcgctagatgcccttctgtcca 1655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 ACATTGTGAACCAAATGACAGAAGCCTGCCTTAACCAGTCGCTAGATGCCCTTCTGTCCA 1620

Qy   1656 gggacttgatcatgaaagaggactatgaacttgttagtaccaagcctacaaggacctcaa 1715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GGGACTTGATCATGAAAGAGGACTATGAACTTGTTAGTACCAAGCCTACAAGGACCTCAA 1680

Qy   1716 aagtcagacaattactagacactactgacatccaaggagaagaatttgccaaagttatag 1775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 AAGTCAGACAATTACTAGACACTACTGACATCCAAGGAGAAGAATTTGCCAAAGTTATAG 1740

Qy   1776 tacaaaaattgaaagataacaaacaaatgggtcttcagccttacccggaaatacttgtgg 1835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 TACAAAAATTGAAAGATAACAAACAAATGGGTCTTCAGCCTTACCCGGAAATACTTGTGG 1800

Qy   1836 tttctagatcaccatctttaaatttacttcaaaataaaagcatgtaagtgactgtttttc 1895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 TTTCTAGATCACCATCTTTAAATTTACTTCAAAATAAAAGCATGTAAGTGACTGTTTTTC 1860

Qy   1896 aagaagaaatgtgtttcataaaaggatatttatatctctgttgctttgactttttttata 1955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 AAGAAGAAATGTGTTTCATAAAAGGATATTTATATCTCTGTTGCTTTGACTTTTTTTATA 1920

Qy   1956 taaaatccgtgagtattaaagcttwawwraargktctttsrktaaatattagtctccctc 2015
          ||||||||||||||||||||||||:|:::||:|:|||||:::||||||||||||||||||
Db   1921 TAAAATCCGTGAGTATTAAAGCTTTATTGAAGGTTCTTTGGGTAAATATTAGTCTCCCTC 1980

Qy   2016 catgacactgcagtattttttttaattaatacaagtaaaaagttgaatttg 2066
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 CATGACACTGCAGTATTTTTTTTAATTAATACAAGTAAAAAGTTGAATTTG 2031
```

```
RESULT    3
AF064824
LOCUS       AF064824      1902 bp    mRNA            PRI       07-JUL-1998
DEFINITION  Homo sapiens CARD-containing ICE associated kinase mRNA, complete
            cds.
ACCESSION   AF064824
VERSION     AF064824.1  GI:3290171
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1902)
  AUTHORS   Thome,M., Hofmann,K., Burns,K., Martinon,F., Bodmer,J.-L.,
            Mattmann,C. and Tschopp,J.
  TITLE     Identification of CARDIAK, a RIP-like kinase that associates with
            caspase-1
  JOURNAL   Curr. Biol. 8 (1998) In press
REFERENCE   2  (bases 1 to 1902)
  AUTHORS   Thome,M., Hofmann,K., Burns,K., Martinon,F., Bodmer,J.-L.,
            Mattmann,C. and Tschopp,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAY-1998) Institute of Biochemistry, University of
            Lausanne, Chemin des Boveresses, 155, Epalinges, CH 1066,
            Switzerland
FEATURES             Location/Qualifiers
     source          1. .1902
```

```
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /db_xref="dbEST:AA161113"
                     /clone_lib="Stratagene library, catalog number 937208"
                     /cell_type="pancreatic adenocarcinoma"
     CDS             243. .1865
                     /note="CARDIAK; RIP-like kinase"
                     /codon_start=1
                     /product="CARD-containing ICE associated kinase"
                     /protein_id="AAC25668.1"
                     /db_xref="GI:3290172"
                     /translation="MNGEAICSALPTIPYHKLADLRYLSRGASGTVSSARHADWRVQV
                     AVKHLHIHTPLLDSERKDVLREAEILHKARFSYILPILGICNEPEFLGIVTEYMPNGS
                     LNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHV
                     KIADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVI
                     TWEVLSRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIPHRARMISLIESGWAQN
                     PDERPSFLKCLIELEPVLRTFEEITFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLN
                     IPVNHGPQEESCGSSQLHENSGSPETSRSLPAPQDNDFLSRKAQDCYFMKLHHCPGNH
                     SWDSTISGSQRAAFCDHKTTPCSSAIINPLSTAGNSERLQPGIAQQWIQSKREDIVNQ
                     MTEACLNQSLDALLSRDLIMKEDYELVSTKPTRTSKVRQLLDTTDIQGEEFAKVIVQK
                     LKDNKQMGLQPYPEILVVSRSPSLNLLQNKSM"
     misc_feature    294. .1124
                     /note="encodes kinase domain"
     misc_feature    1125. .1544
                     /note="encodes intermediate domain"
     misc_feature    1545. .1811
                     /note="encodes CARD motif"
BASE COUNT      582 a    432 c    419 g    468 t      1 others
ORIGIN


  Query Match            86.5%;  Score 1815.2;  DB 88;  Length 1902;
  Best Local Similarity  99.6%;  Pred. No. 0;
  Matches 1861;  Conservative    0;  Mismatches    4;  Indels    4;  Gaps    4;

Qy         40 ccagtctctagaaaagaagtcagctctggttcggaga-agcagcggctggcg-tgggcca 97
              ||||||||||||||||||||||||||||||||||||| |||||||||||||| |||||||
Db         19 CCAGTCTCTAGAAAAGAAGTCAGCTCTGGTTCGGAGATAGCAGCGGCTGGCGTTGGGCCA 78

Qy         98 tccggggaatgggcgccctcgtgacct-agtgttgcggggcaaaaagggtc-ttgccggc 155
              ||||||||||||||||||||||||||| ||||||||||||||||||||||| ||||||||
Db         79 TCCGGGGAATGGGCGCCCTCGTGACCTGAGTGTTGCGGGGCAAAAAGGGTCTTTGCCGGC 138

Qy        156 ctcgctcgtgcaggggcgtatctgggcgcctgagcgcggcgtgggagccttgggagccgc 215
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        139 CTCGCTCGTGCAGGGGCGTATCTGGNCGCCTGAGCGCGGCGTGGGAGCCTTGGGAGCCGC 198

Qy        216 cgcagcagggggcacacccggaaccggcctgagcgcccgggaccatgaacggggaggcca 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 CGCAGCAGGGGGCACACCCGGAACCGGCCTGAGCGCCCGGGACCATGAACGGGGAGGCCA 258

Qy        276 tctgcagcgccctgcccaccattccctaccacaaactcgccgacctgcgctacctgagcc 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 TCTGCAGCGCCCTGCCCACCATTCCCTACCACAAACTCGCCGACCTGCGCTACCTGAGCC 318

Qy        336 gcggcgcctctggcactgtgtcgtccgcccgccacgcagactggcgcgtccaggtggccg 395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 GCGGCGCCTCTGGCACTGTGTCGTCCGCCCGCCACGCAGACTGGCGCGTCCAGGTGGCCG 378

Qy        396 tgaagcacctgcacatccacactccgctgctcgacagtgaaagaaaggatgtttaagag 455
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        379 TGAAGCACCTGCACATCCACACTCCGCTGCTCGACAGTGAAAGAAAGGATGTCTTAAGAG 438

Qy        456 aagctgaaatttacacaaagctagatttagttacatttttccaatttgggaatttgca 515
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        439 AAGCTGAAATTTACACAAAGCTAGATTTAGTTACATTCTTCCAATTTTGGGAATTTGCA 498

Qy        516 atgagcctgaatttttgggaatagttactgaatacatgccaaatggatcattaaatgaac 575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 ATGAGCCTGAATTTTTGGGAATAGTTACTGAATACATGCCAAATGGATCATTAAATGAAC 558

Qy        576 tcctacataggaaaactgaatatcctgatgttgcttggccattgagatttcgcatcctgc 635
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 TCCTACATAGGAAAACTGAATATCCTGATGTTGCTTGGCCATTGAGATTTCGCATCCTGC 618

Qy        636 atgaaattgcccttggtgtaaattacctgcacaatatgactcctcctttacttcatcatg 695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        619 ATGAAATTGCCCTTGGTGTAAATTACCTGCACAATATGACTCCTCCTTTACTTCATCATG 678

Qy        696 acttgaagactcagaatatcttattggacaatgaatttcatgttaagattgcagattttg 755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        679 ACTTGAAGACTCAGAATATCTTATTGGACAATGAATTTCATGTTAAGATTGCAGATTTTG 738

Qy        756 gtttatcaaagtggcgcatgatgtccctctcacagtcacgaagtagcaaatctgcaccag 815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        739 GTTTATCAAAGTGGCGCATGATGTCCCTCTCACAGTCACGAAGTAGCAAATCTGCACCAG 798

Qy        816 aaggagggacaattatttatatgccacctgaaaactatgaacctggacaaaaatcaaggg 875
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        799 AAGGAGGGACAATTATCTATATGCCACCTGAAAACTATGAACCTGGACAAAAATCAAGGG 858

Qy        876 ccagtatcaagcacgatatatatagctatgcagttatcacatgggaagtgttatccagaa 935
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        859 CCAGTATCAAGCACGATATATATAGCTATGCAGTTATCACATGGGAAGTGTTATCCAGAA 918

Qy        936 aacagcctttgaagatgtcaccaatcctttgcagataatgtatagtgtgtcacaaggac 995
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        919 AACAGCCTTTTGAAGATGTCACCAATCCTTTGCAGATAATGTATAGTGTGTCACAAGGAC 978

Qy        996 atcgacctgttattaatgaagaaagtttgccatatgatatacctcaccgagcacgtatga 1055
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        979 ATCGACCTGTTATTAATGAAGAAAGTTTGCCATATGATATACCTCACCGAGCACGTATGA 1038

Qy       1056 tctctctaatagaaagtggatgggcacaaaatccagatgaaagaccatctttcttaaaat 1115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1039 TCTCTCTAATAGAAAGTGGATGGGCACAAAATCCAGATGAAAGACCATCTTTCTTAAAAT 1098

Qy       1116 gtttaatagaacttgaaccagttttgagaacatttgaagagataacttttcttgaagctg 1175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1099 GTTTAATAGAACTTGAACCAGTTTTGAGAACATTTGAAGAGATAACTTTTCTTGAAGCTG 1158

Qy       1176 ttattcagctaaagaaaacaaagttacagagtgtttcaagtgccattcacctatgtgaca 1235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1159 TTATTCAGCTAAAGAAAACAAAGTTACAGAGTGTTTCAAGTGCCATTCACCTATGTGACA 1218

Qy       1236 agaagaaaatggaattatctctgaacatacctgtaaatcatggtccacaagaggaatcat 1295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1219 AGAAGAAAATGGAATTATCTCTGAACATACCTGTAAATCATGGTCCACAAGAGGAATCAT 1278

Qy       1296 gtggatcctctcagctccatgaaaatagtggttctcctgaaacttcaaggtccctgccag 1355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1279 GTGGATCCTCTCAGCTCCATGAAAATAGTGGTTCTCCTGAAACTTCAAGGTCCCTGCCAG 1338

Qy       1356 ctcctcaagacaatgatttttatctagaaaagctcaagactgttattttatgaagctgc 1415
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1339 CTCCTCAAGACAATGATTTTTATCTAGAAAAGCTCAAGACTGTTATTTTATGAAGCTGC 1398

Qy       1416 atcactgtcctggaaatcacagttgggatagcaccatttctggatctcaaagggctgcat 1475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1399 ATCACTGTCCTGGAAATCACAGTTGGGATAGCACCATTTCTGGATCTCAAAGGGCTGCAT 1458

Qy       1476 tctgtgatcacaagaccactccatgctcttcagcaataataaatccactctcaactgcag 1535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1459 TCTGTGATCACAAGACCACTCCATGCTCTTCAGCAATAATAAATCCACTCTCAACTGCAG 1518

Qy       1536 gaaactcagaacgtctgcagcctggtatagcccagcagtggatccagagcaaaagggaag 1595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1519 GAAACTCAGAACGTCTGCAGCCTGGTATAGCCCAGCAGTGGATCCAGAGCAAAAGGGAAG 1578

Qy       1596 acattgtgaaccaaatgacagaagcctgccttaaccagtcgctagatgcccttctgtcca 1655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1579 ACATTGTGAACCAAATGACAGAAGCCTGCCTTAACCAGTCGCTAGATGCCCTTCTGTCCA 1638

Qy       1656 gggacttgatcatgaaagaggactatgaacttgttagtaccaagcctacaaggacctcaa 1715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1639 GGGACTTGATCATGAAAGAGGACTATGAACTTGTTAGTACCAAGCCTACAAGGACCTCAA 1698
```

```
Qy    1716 aagtcagacaattactagacactactgacatccaaggagaagaatttgccaaagttatag 1775
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1699 AAGTCAGACAATTACTAGACACTACTGACATCCAAGGAGAAGAATTTGCCAAAGTTATAG 1758

Qy    1776 tacaaaaattgaaagataacaaacaaatgggtcttcagccttacccggaaatacttgtgg 1835
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1759 TACAAAAATTGAAAGATAACAAACAAATGGGTCTTCAGCCTTACCCGGAAATACTTGTGG 1818

Qy    1836 tttctagatcaccatctttaaatttacttcaaaataaaagcatgtaagtgactgtttttc 1895
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1819 TTTCTAGATCACCATCTTTAAATTTACTTCAAAATAAAAGCATGTAAGTGACTGTTTTTC 1878

Qy    1896 aagaagaaa 1904
           |||||||||
Db    1879 AAGAAGAAA 1887


RESULT   4
AF078530
LOCUS       AF078530     1623 bp    mRNA            PRI       28-JUL-1998
DEFINITION  Homo sapiens receptor interacting protein 2 (RIP2) mRNA, complete
            cds.
ACCESSION   AF078530
VERSION     AF078530.1  GI:3342909
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1623)
  AUTHORS   McCarthy,J.V., Ni,J. and Dixit,V.M.
  TITLE     RIP2 is a novel NF-kappaB-activating and cell death-inducing kinase
  JOURNAL   J. Biol. Chem. 273 (27), 16968-16975 (1998)
  MEDLINE   98307936
REFERENCE   2  (bases 1 to 1623)
  AUTHORS   McCarthy,J.V., Ni,J. and Dixit,V.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUL-1998) Molecular Oncology, Genentech Inc, 1 DNA
            Way, South San Francisco, CA 94080, USA
FEATURES             Location/Qualifiers
     source          1..1623
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            1..1623
                     /gene="RIP2"
     CDS             1..1623
                     /gene="RIP2"
                     /function="activtes NF-kappaB"
                     /function="induces cell death"
                     /note="RIP2; serine/threonine kinase homolog; contains
                     CARD motif"
                     /codon_start=1
                     /product="receptor interacting protein 2"
                     /protein_id="AAC27722.1"
                     /db_xref="GI:3342910"
                     /translation="MNGEAICSALPTIPYHKLADLRYLSRGASGTVSSARHADWRVQV
                     AVKHLHIHTPLLDSERKDVLREAEILHKARFSYILPILGICNEPEFLGIVTEYMPNGS
                     LNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHV
                     KIADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVI
                     TWEVLSRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIPHRARMISLIESGWAQN
                     PDERPSFLKCLIELEPVLRTFEEITFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLN
                     IPVNHGPQEESCGSSQLHENSGSPETSRSLPAPQDNDFLSRKAQDCYFMKLHHCPGNH
                     SWDSTISGSQRAAFCDHKTTPCSSAIINPLSTAGNSERLQPGIAQQWIQSKREDIVNQ
                     MTEACLNQSLDALLSRDLIMKEDYELVSTKPTRTSKVRQLLDTTDIQGEEFAKVIVQK
                     LKDNKQMGLQPYPEILVVSRSPSLNLLQNKSM"
BASE COUNT      519 a    362 c    320 g    422 t
ORIGIN


  Query Match             77.1%;  Score 1616.6;  DB 88;  Length 1623;
  Best Local Similarity   99.8%;  Pred. No. 1.8e-277;
  Matches 1619;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy     260 atgaacggggaggccatctgcagcgccctgcccaccattccctaccacaaactcgccgac 319
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGAACGGGGAGGCCATCTGCAGCGCCCTGCCCACCATTCCCTACCACAAACTCGCCGAC 60

Qy     320 ctgcgctacctgagccgcggcgcctctggcactgtgtcgtccgcccgccacgcagactgg 379
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CTGCGCTACCTGAGCCGCGGCGCCTCTGGCACTGTGTCGTCCGCCCGCCACGCAGACTGG 120

Qy     380 cgcgtccaggtggccgtgaagcacctgcacatccacactccgctgctcgacagtgaaaga 439
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CGCGTCCAGGTGGCCGTGAAGCACCTGCACATCCACACTCCGCTGCTCGACAGTGAAAGA 180

Qy     440 aaggatgttttaagagaagctgaaattttacacaaagctagatttagttacatttttcca 499
           |||||||| |||||||||||||||||||||||||||||||||||||||||||| |||||
Db     181 AAGGATGTCTTAAGAGAAGCTGAAATTTTACACAAAGCTAGATTTAGTTACATTCTTCCA 240

Qy     500 atttgggaatttgcaatgagcctgaatttttgggaatagttactgaatacatgccaaat 559
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ATTTTGGGAATTTGCAATGAGCCTGAATTTTTGGGAATAGTTACTGAATACATGCCAAAT 300

Qy     560 ggatcattaaatgaactcctacataggaaaactgaatatcctgatgttgcttggccattg 619
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GGATCATTAAATGAACTCCTACATAGGAAAACTGAATATCCTGATGTTGCTTGGCCATTG 360

Qy     620 agatttcgcatcctgcatgaaattgcccttggtgtaaattacctgcacaatatgactcct 679
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AGATTTCGCATCCTGCATGAAATTGCCCTTGGTGTAAATTACCTGCACAATATGACTCCT 420

Qy     680 cctttacttcatcatgacttgaagactcagaatatcttattggacaatgaatttcatgtt 739
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CCTTTACTTCATCATGACTTGAAGACTCAGAATATCTTATTGGACAATGAATTTCATGTT 480

Qy     740 aagattgcagattttggtttatcaaagtggcgcatgatgtccctctcacagtcacgaagt 799
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AAGATTGCAGATTTTGGTTTATCAAAGTGGCGCATGATGTCCCTCTCACAGTCACGAAGT 540

Qy     800 agcaaatctgcaccagaaggagggacaattatttatatgccacctgaaaactatgaacct 859
           |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db     541 AGCAAATCTGCACCAGAAGGAGGGACAATTATCTATATGCCACCTGAAAACTATGAACCT 600

Qy     860 ggacaaaaatcaagggccagtatcaagcacgatatatatagctatgcagttatcacatgg 919
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GGACAAAAATCAAGGGCCAGTATCAAGCACGATATATATAGCTATGCAGTTATCACATGG 660

Qy     920 gaagtgttatccagaaaacagccttttgaagatgtcaccaatcctttgcagataatgtat 979
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GAAGTGTTATCCAGAAAACAGCCTTTTGAAGATGTCACCAATCCTTTGCAGATAATGTAT 720

Qy     980 agtgtgtcacaaggacatcgacctgttattaatgaagaaagtttgccatatgatatacct 1039
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 AGTGTGTCACAAGGACATCGACCTGTTATTAATGAAGAAAGTTTGCCATATGATATACCT 780

Qy    1040 caccgagcacgtatgatctctctaatagaaagtggatgggcacaaaatccagatgaaaga 1099
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CACCGAGCACGTATGATCTCTCTAATAGAAAGTGGATGGGCACAAAATCCAGATGAAAGA 840

Qy    1100 ccatctttcttaaaatgtttaatagaacttgaaccagttttgagaacatttgaagagata 1159
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CCATCTTTCTTAAAATGTTTAATAGAACTTGAACCAGTTTTGAGAACATTTGAAGAGATA 900

Qy    1160 actttttcttgaagctgttattcagctaaagaaaacaaagttacagagtgtttcaagtgcc 1219
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ACTTTTCTTGAAGCTGTTATTCAGCTAAAGAAAACAAAGTTACAGAGTGTTTCAAGTGCC 960

Qy    1220 attcacctatgtgacaagaagaaaatggaattatctctgaacatacctgtaaatcatggt 1279
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ATTCACCTATGTGACAAGAAGAAAATGGAATTATCTCTGAACATACCTGTAAATCATGGT 1020

Qy    1280 ccacaagaggaatcatgtggatcctctcagctccatgaaaatagtggttctcctgaaact 1339
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 CCACAAGAGGAATCATGTGGATCCTCTCAGCTCCATGAAAATAGTGGTTCTCCTGAAACT 1080
```

```
Qy    1340 tcaaggtccctgccagctcctcaagacaatgatttttatctagaaaagctcaagactgt 1399
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 TCAAGGTCCCTGCCAGCTCCTCAAGACAATGATTTTTATCTAGAAAAGCTCAAGACTGT 1140

Qy    1400 tatttatgaagctgcatcactgtcctggaaatcacagttgggatagcaccatttctgga 1459
           ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    1141 TATTTATGAAGCTGCATCACTGTCCTGGAAATCACAGTTGGGACAGCACCATTTCTGGA 1200

Qy    1460 tctcaaagggctgcattctgtgatcacaagaccactccatgctcttcagcaataataaat 1519
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TCTCAAAGGGCTGCATTCTGTGATCACAAGACCACTCCATGCTCTTCAGCAATAATAAAT 1260

Qy    1520 ccactctcaactgcaggaaactcagaacgtctgcagcctggtatagcccagcagtggatc 1579
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 CCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAGCCTGGTATAGCCCAGCAGTGGATC 1320

Qy    1580 cagagcaaaagggaagacattgtgaaccaaatgacagaagcctgccttaaccagtcgcta 1639
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 CAGAGCAAAAGGGAAGACATTGTGAACCAAATGACAGAAGCCTGCCTTAACCAGTCGCTA 1380

Qy    1640 gatgcccttctgtccagggacttgatcatgaaagaggactatgaacttgttagtaccaag 1699
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 GATGCCCTTCTGTCCAGGGACTTGATCATGAAAGAGGACTATGAACTTGTTAGTACCAAG 1440

Qy    1700 cctacaaggacctcaaaagtcagacaattactagacactactgacatccaaggagaagaa 1759
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 CCTACAAGGACCTCAAAAGTCAGACAATTACTAGACACTACTGACATCCAAGGAGAAGAA 1500

Qy    1760 tttgccaaagttatagtacaaaaattgaaagataacaaacaaatgggtcttcagccttac 1819
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 TTTGCCAAAGTTATAGTACAAAAATTGAAAGATAACAAACAAATGGGTCTTCAGCCTTAC 1560

Qy    1820 ccggaaatacttgtggtttctagatcaccatctttaaatttacttcaaaataaaagcatg 1879
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 CCGGAAATACTTGTGGTTTCTAGATCACCATCTTTAAATTTACTTCAAAATAAAAGCATG 1620

Qy    1880 taa 1882
           |||
Db    1621 TAA 1623


RESULT   5
AC004003
LOCUS       AC004003   116650 bp    DNA             PRI       21-DEC-1999
DEFINITION  Homo sapiens BAC clone CTA-437L15 from 8q21, complete sequence.
ACCESSION   AC004003
VERSION     AC004003.1  GI:2772557
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 116650)
  AUTHORS   Ozersky,P., Holmes,A. and Broy,M.
  TITLE     The sequence of Homo sapiens BAC clone CTA-437L15
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 116650)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-1998) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   3  (bases 1 to 116650)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JUN-1998) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   4  (bases 1 to 116650)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-DEC-1999) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     -------------- Genome Center
                Center: Washington University Genome Sequencing Center
```

```
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_RG437L15
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded or sequenced with an alternate
            chemistry; an attempt was made to resolve all sequencing problems,
            such as compressions and repeats; all regions were covered by
            sequence from more than one subclone; and the assembly was
            confirmed by restriction digest.

            MAPPING INFORMATION:
            This chromosome 8 clone was provided by Dr. Patrick Concannon
            (patcon@vmmc.org) at the Virginia Mason Research Institute.

            SOURCE INFORMATION:
            Clone CTA-437L15 is from a release of the human BAC library
            CITB-HS-A.  The library contains cloned DNA from human sperm.  See:
             Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J.
            Kim et al., Genomics 34:213-8 (1996).  The clone is available from
            Research Genetics, Inc. (http://www.resgen.com).
            VECTOR:  pBeloBAC11
            Selection:  chloramphenicol

            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is CTA-237G1. The actual start of
            this clone is at base position 1 of CTA-437L15; actual end is at
            116650 of CTA-437L15.

            This clone contains STS HS275YF1 (NID:g1051703).
FEATURES             Location/Qualifiers
     source          1. .116650
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8q21"
                     /clone="CTA-437L15"
                     /clone_lib="CITB-HS-A"
     repeat_region   150. .530
                     /rpt_family="L1"
     repeat_region   498. .649
                     /rpt_family="L1"
     repeat_region   673. .1096
                     /rpt_family="MaLR"
     repeat_region   3188. .3348
                     /rpt_family="L1"
     repeat_region   3605. .3726
                     /rpt_family="(TA)n"
     repeat_region   3845. .4563
                     /rpt_family="L1"
     repeat_region   5641. .6075
                     /rpt_family="Retroviral"
     repeat_region   6133. .6433
                     /rpt_family="Alu"
     misc_feature    7110. .8040
                     /note="CpG_island (%GC=69.1, o/e=0.80, #CpGs=93)"
     gene            7872. .40218
                     /gene="WUGSC:H_RG437L15.1"
     CDS             join(7872. .8044,12642. .12795,15154. .15309,19584. .19741,
                     21552. .21601,22478. .22639,29890. .29975,33855. .33944,
                     36395. .36488,39123. .39284,39881. .40218)
                     /gene="WUGSC:H_RG437L15.1"
                     /note="serine/threonine kinase RICK; match to protein
                     AF027706 (PID:g3123887) and mRNA AF027706 (NID:g3123886);
                     H_RG437L15.1"
```

```
                    /codon_start=1
                    /evidence=not_experimental
                    /protein_id="AAC24561.1"
                    /db_xref="GI:3264574"
                    /translation="MNGEAICSALPTIPYHKLADLRYLSRGASGTVSSARHADWRVQV
                    AVKHLHIHTPLLDSERKDVLREAEILHKARFSYILPILGICNEPEFLGIVTEYMPNGS
                    LNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHV
                    KIADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVI
                    TWEVLSRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIPHRARMISLIESGWAQN
                    PDERPSFLKCLIELEPVLRTFEEITFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLN
                    IPVNHGPQEESCGSSQLHENSGSPETSRSLPAPQDNDFLSRKAQDCYFMKLHHCPGNH
                    SWDSTISGSQRAAFCDHKTTPCSSAIINPLSTAGNSERLQPGIAQQWIQSKREDIVNQ
                    MTEACLNQSLDALLSRDLIMKEDYELVSTKPTRTSKVRQLLDTTDIQGEEFAKVIVQK
                    LKDNKQMGLQPYPEILVVSRSPSLNLLQNKSM"
     repeat_region  8463. .8770
                    /rpt_family="MER1_type"
     repeat_region  10435. .10540
                    /rpt_family="MIR"
     repeat_region  11433. .11538
                    /rpt_family="Alu"
     repeat_region  11790. .12003
                    /rpt_family="Alu"
     repeat_region  12466. .12557
                    /rpt_family="MIR"
     repeat_region  15579. .16292
                    /rpt_family="L1"
     repeat_region  16292. .16429
                    /rpt_family="L1"
     repeat_region  16460. .16991
                    /rpt_family="L1"
     repeat_region  22863. .23483
                    /rpt_family="L2"
     repeat_region  23575. .23750
                    /rpt_family="L1"
     repeat_region  23761. .23924
                    /rpt_family="Alu"
     repeat_region  23929. .24439
                    /rpt_family="L1"
     repeat_region  24471. .24960
                    /rpt_family="L2"
     repeat_region  24967. .25285
                    /rpt_family="Alu"
     repeat_region  25337. .25671
                    /rpt_family="L2"
     repeat_region  26151. .26280
                    /rpt_family="MER1_type"
     repeat_region  26268. .26500
                    /rpt_family="MER1_type"
     repeat_region  26611. .27000
                    /rpt_family="L1"
     repeat_region  28128. .29109
                    /rpt_family="L1"
     repeat_region  29128. .29434
                    /rpt_family="Alu"
     repeat_region  29435. .29752
                    /rpt_family="L1"
     repeat_region  30205. .30341
                    /rpt_family="Alu"
     repeat_region  31170. .31344
                    /rpt_family="Alu"
     repeat_region  31466. .31688
                    /rpt_family="MER2_type"
     repeat_region  31761. .31955
                    /rpt_family="Alu"
     repeat_region  31956. .32225
                    /rpt_family="MER2_type"
     repeat_region  32329. .32460
                    /rpt_family="Alu"
     repeat_region  33155. .33281
                    /rpt_family="MIR"
     repeat_region  33380. .33437
                    /rpt_family="MIR"
     repeat_region  34433. .34514
                    /rpt_family="MER2_type"
     repeat_region  34525. .34939
                    /rpt_family="MaLR"
     repeat_region  34999. .35256
                    /rpt_family="Alu"
     repeat_region  35388. .35477
                    /rpt_family="MER2_type"
     repeat_region  35994. .36286
                    /rpt_family="Alu"
     repeat_region  36686. .36985
                    /rpt_family="Alu"
     repeat_region  37488. .37673
                    /rpt_family="L2"
     repeat_region  37862. .38161
                    /rpt_family="Alu"
     repeat_region  38297. .38599
                    /rpt_family="Alu"
     repeat_region  41954. .42006
                    /rpt_family="L2"
     repeat_region  42596. .42777
                    /rpt_family="MER1_type"
     repeat_region  43357. .43658
                    /rpt_family="Alu"
     repeat_region  44088. .44386

  Query Match             24.6%;  Score 516.6;  DB 85;  Length 116650;
  Best Local Similarity   97.7%;  Pred. No. 3.7e-82;
  Matches  513;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1541 tcagaacgtctgcagcctggtatagcccagcagtggatccagagcaaaagggaagacatt 1600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      39877 TCAGAACGTCTGCAGCCTGGTATAGCCCAGCAGTGGATCCAGAGCAAAAGGGAAGACATT 39936

Qy       1601 gtgaaccaaatgacagaagcctgccttaaccagtcgctagatgcccttctgtccagggac 1660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      39937 GTGAACCAAATGACAGAAGCCTGCCTTAACCAGTCGCTAGATGCCCTTCTGTCCAGGGAC 39996

Qy       1661 ttgatcatgaaagaggactatgaacttgttagtaccaagcctacaaggacctcaaaagtc 1720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      39997 TTGATCATGAAAGAGGACTATGAACTTGTTAGTACCAAGCCTACAAGGACCTCAAAAGTC 40056

Qy       1721 agacaattactagacactactgacatccaaggagaagaatttgccaaagttatagtacaa 1780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      40057 AGACAATTACTAGACACTACTGACATCCAAGGAGAAGAATTTGCCAAAGTTATAGTACAA 40116

Qy       1781 aaattgaaagataacaaacaaatgggtcttcagccttacccggaaatacttgtggtttct 1840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      40117 AAATTGAAAGATAACAAACAAATGGGTCTTCAGCCTTACCCGGAAATACTTGTGGTTTCT 40176

Qy       1841 agatcaccatctttaaatttacttcaaaataaaagcatgtaagtgactgtttttcaagaa 1900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      40177 AGATCACCATCTTTAAATTTACTTCAAAATAAAAGCATGTAAGTGACTGTTTTTCAAGAA 40236

Qy       1901 gaaatgtgtttcataaaaggatatttatatctctgttgctttgactttttttatataaaa 1960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      40237 GAAATGTGTTTCATAAAAGGATATTTATATCTCTGTTGCTTTGACTTTTTTTATATAAAA 40296

Qy       1961 tccgtgagtattaaagcttwawwraargktctttsrktaaatattagtctccctccatga 2020
              |||||||||||||||||||:|:::||:|:|||||:::|||||||||||||||||||||||
Db      40297 TCCGTGAGTATTAAAGCTTTATTGAAGGTTCTTTGGTAAATATTAGTCTCCCTCCATGA 40356

Qy       2021 cactgcagtatttttttttaattaatacaagtaaaagttgaattt 2065
              |||||||||||||||||||||||||||||||||||||||  |  ||
Db      40357 CACTGCAGTATTTTTTTTTAATTAATACAAGTAAAAGTTTGAATT 40401


RESULT   6
AF117829
LOCUS       AF117829   320250 bp    DNA             PRI       13-JAN-1999
DEFINITION  Homo sapiens 8q21.3: RICK gene, complete sequence.
ACCESSION   AF117829
VERSION     AF117829.1  GI:4151947
KEYWORDS    HTG.
```

```
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 320250)
  AUTHORS   Platzer,M. and Varon,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-DEC-1998) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
COMMENT     This sequence is part of a larger genomic contig. The start of this
            sequence is directed towards the centromere. The end of this
            sequence (320001. .320250) overlaps with the start of the
            neighbouring Acc_number AF049895 (1. .250).
FEATURES             Location/Qualifiers
     source          1. .320250
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8q21.3"
     source          1. .137184
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="P64M4"
     repeat_region   complement(572. .704)
                     /rpt_family="MIR"
     repeat_region   705. .870
                     /rpt_family="MIR"
     repeat_region   complement(930. .985)
                     /rpt_family="MIR"
     repeat_region   2222. .2456
                     /rpt_family="L2"
     repeat_region   2538. .2661
                     /rpt_family="L2"
     repeat_region   complement(2695. .2845)
                     /rpt_family="FAM"
     repeat_region   2913. .3485
                     /rpt_family="L2"
     repeat_region   3559. .3679
                     /rpt_family="L2"
     repeat_region   3776. .3930
                     /rpt_family="MIR"
     repeat_region   complement(4035. .4247)
                     /rpt_family="MIR"
     repeat_region   complement(5211. .5655)
                     /rpt_family="L2"
     repeat_region   5997. .6100
                     /rpt_family="L2"
     repeat_region   6800. .6850
                     /rpt_family="(TAGA)n"
     repeat_region   8389. .8709
                     /rpt_family="L2"
     repeat_region   complement(9413. .9761)
                     /rpt_family="MLT1A1"
     repeat_region   10399. .10473
                     /rpt_family="L2"
     repeat_region   11341. .11507
                     /rpt_family="MIR"
     repeat_region   complement(11525. .11570)
                     /rpt_family="(TAA)n"
     repeat_region   11931. .11985
                     /rpt_family="(CATG)n"
     repeat_region   complement(12051. .12149)
                     /rpt_family="MER53"
     repeat_region   complement(12351. .12760)
                     /rpt_family="MSTA"
     repeat_region   14544. .14734
                     /rpt_family="MIR"
     repeat_region   complement(15740. .15847)
                     /rpt_family="MIR"
     repeat_region   16491. .16558
                     /rpt_family="MIR"
     repeat_region   18103. .18162
                     /rpt_family="MIR"
     repeat_region   18343. .18639
                     /rpt_family="AluSx"
     repeat_region   19005. .19075
                     /rpt_family="MER57-internal"
     repeat_region   complement(19478. .19560)
                     /rpt_family="MIR"
     repeat_region   complement(23416. .23710)
                     /rpt_family="AluSg"
     repeat_region   complement(24833. .25025)
                     /rpt_family="MIR"
     repeat_region   complement(25333. .25392)
                     /rpt_family="MER81"
     repeat_region   25969. .26053
                     /rpt_family="L1MB8"
     repeat_region   26167. .26302
                     /rpt_family="L1M4"
     repeat_region   complement(26512. .26684)
                     /rpt_family="AluJo"
     repeat_region   26685. .26712
                     /rpt_family="(TA)n"
     repeat_region   complement(26713. .26841)
                     /rpt_family="AluJo"
     repeat_region   complement(27456. .27620)
                     /rpt_family="MER5A"
     repeat_region   complement(27724. .28020)
                     /rpt_family="AluY"
     repeat_region   28147. .28283
                     /rpt_family="AluY"
     repeat_region   28284. .28311
                     /rpt_family="(TA)n"
     repeat_region   28312. .28476
                     /rpt_family="AluY"
     repeat_region   29222. .29518
                     /rpt_family="L1M2"
     repeat_region   29454. .32583
                     /rpt_family="L1MA6"
     repeat_region   32636. .33403
                     /rpt_family="L1MA6"
     repeat_region   33404. .33705
                     /rpt_family="AluSc"
     repeat_region   33706. .33960
                     /rpt_family="L1MA6"
     repeat_region   complement(34346. .34614)
                     /rpt_family="MER8"
     repeat_region   34786. .34873
                     /rpt_family="MIR"
     repeat_region   34839. .35057
                     /rpt_family="L2"
     repeat_region   complement(35173. .35382)
                     /rpt_family="AluYa5"
     repeat_region   35577. .35746
                     /rpt_family="MLT1D"
     repeat_region   complement(35747. .36046)
                     /rpt_family="AluY"
     repeat_region   36047. .36323
                     /rpt_family="MLT1D"
     repeat_region   complement(36593. .36761)
                     /rpt_family="MIR"
     repeat_region   complement(37858. .37899)
                     /rpt_family="L2"
     repeat_region   38652. .38895
                     /rpt_family="L1M2"
     repeat_region   38959. .39811
                     /rpt_family="L1M3e"
     repeat_region   39885. .39989
                     /rpt_family="L1M2"
     repeat_region   40242. .43133
                     /rpt_family="L1M2"
     repeat_region   43134. .43499
                     /rpt_family="MLT1A1"
     repeat_region   43500. .44075
                     /rpt_family="L1M2"
     repeat_region   44076. .44391
```

```
                     /rpt_family="AluJo"
     repeat_region   44392. .45758
                     /rpt_family="L1MA4A"
     repeat_region   complement(45759. .46216)
                     /rpt_family="L1PA7"
     repeat_region   46217. .47676
                     /rpt_family="L1MA4A"
     repeat_region   47693. .47955
                     /rpt_family="L1MA5A"
     repeat_region   48227. .48421
                     /rpt_family="L2"
     repeat_region   complement(49170. .49277)
                     /rpt_family="MIR"
     repeat_region   complement(49409. .50232)
                     /rpt_family="L1PA2"
     repeat_region   50233. .50476
                     /rpt_family="L1PA4"
     repeat_region   complement(50588. .50644)
                     /rpt_family="MIR"
     repeat_region   complement(51729. .52082)
                     /rpt_family="THE1C"
     repeat_region   complement(52083. .53629)
                     /rpt_family="THE1C-internal"
     repeat_region   complement(53632. .53992)
                     /rpt_family="THE1C"
     repeat_region   54101. .54351
                     /rpt_family="L2"
     repeat_region   54433. .54822
                     /rpt_family="L2"
     repeat_region   54823. .55177
                     /rpt_family="MLT1A2"
     repeat_region   55178. .55291
                     /rpt_family="L2"
     repeat_region   complement(55316. .55371)
                     /rpt_family="L2"
     repeat_region   55545. .55878
                     /rpt_family="L2"
     repeat_region   57436. .57481
                     /rpt_family="(TTAAA)n"
     repeat_region   60524. .60721
                     /rpt_family="L1ME"
     repeat_region   complement(61481. .61865)
                     /rpt_family="MSTA"
     repeat_region   62004. .62295
                     /rpt_family="MLT1E"
     repeat_region   62427. .67614
                     /rpt_family="L1PA2"
     source          62833. .145579
                     /organism="Homo sapiens"

  Query Match             24.6%;  Score 516.6;  DB 88;  Length 320250;
  Best Local Similarity   97.7%;  Pred. No. 3.6e-82;
  Matches  513;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy   1541 tcagaacgtctgcagcctggtatagcccagcagtggatccagagcaaaagggaagacatt 1600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 179446 TCAGAACGTCTGCAGCCTGGTATAGCCCAGCAGTGGATCCAGAGCAAAAGGGAAGACATT 179505

Qy   1601 gtgaaccaaatgacagaagcctgccttaaccagtcgctagatgcccttctgtccagggac 1660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 179506 GTGAACCAAATGACAGAAGCCTGCCTTAACCAGTCGCTAGATGCCCTTCTGTCCAGGGAC 179565

Qy   1661 ttgatcatgaaagaggactatgaacttgttagtaccaagcctacaaggacctcaaaagtc 1720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 179566 TTGATCATGAAAGAGGACTATGAACTTGTTAGTACCAAGCCTACAAGGACCTCAAAAGTC 179625

Qy   1721 agacaattactagacactactgacatccaaggagaagaatttgccaaagttatagtacaa 1780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 179626 AGACAATTACTAGACACTACTGACATCCAAGGAGAAGAATTTGCCAAAGTTATAGTACAA 179685

Qy   1781 aaattgaaagataacaaacaaatgggtcttcagccttacccggaaatacttgtggtttct 1840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 179686 AAATTGAAAGATAACAAACAAATGGGTCTTCAGCCTTACCCGGAAATACTTGTGGTTTCT 179745
```

```
Qy   1841 agatcaccatctttaaatttacttcaaaataaaagcatgtaagtgactgtttttcaagaa 1900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 179746 AGATCACCATCTTTAAATTTACTTCAAAATAAAAGCATGTAAGTGACTGTTTTTCAAGAA 179805

Qy   1901 gaaatgtgtttcataaaaggatatttatatctctgttgctttgactttttttatataaaa 1960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 179806 GAAATGTGTTTCATAAAAGGATATTTATATCTCTGTTGCTTTGACTTTTTTTATATAAAA 179865

Qy   1961 tccgtgagtattaaagcttwawwraargktctttsrktaaatattagtctccctccatga 2020
          |||||||||||||||||||:|:::||:|:|||||:::|||||||||||||||||||||||
Db 179866 TCCGTGAGTATTAAAGCTTTATTGAAGGTTCTTTGGGTAAATATTAGTCTCCCTCCATGA 179925

Qy   2021 cactgcagtatttttttaattaatacaagtaaaaagttgaattt 2065
          |||||||||||||||||||||||||||||||||||||||  | ||
Db 179926 CACTGCAGTATTTTTTTAATTAATACAAGTAAAAGTTTGAATT 179970


RESULT   7
AR044139
LOCUS       AR044139       210 bp    DNA             PAT       29-SEP-1999
DEFINITION  Sequence 31 from patent US 5817479.
ACCESSION   AR044139
VERSION     AR044139.1  GI:5965604
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 210)
  AUTHORS   Au-Young,J., Bandman,O., Hawkins,P.R. and Wilde,C.G.
  TITLE     Human kinase homologs
  JOURNAL   Patent: US 5817479-A 31 06-OCT-1998;
FEATURES             Location/Qualifiers
     source          1. .210
                     /organism="unknown"
BASE COUNT      74 a     40 c     41 g     55 t
ORIGIN


  Query Match             10.0%;  Score 210;  DB 9;  Length 210;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-27;
  Matches  210;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    844 tgaaaactatgaacctggacaaaaatcaagggccagtatcaagcacgatatatatagcta 903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TGAAAACTATGAACCTGGACAAAAATCAAGGGCCAGTATCAAGCACGATATATATAGCTA 60

Qy    904 tgcagttatcacatgggaagtgttatccagaaaacagccttttgaagatgtcaccaatcc 963
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TGCAGTTATCACATGGGAAGTGTTATCCAGAAAACAGCCTTTTGAAGATGTCACCAATCC 120

Qy    964 tttgcagataatgtatagtgtgtcacaaggacatcgacctgttattaatgaagaaagttt 1023
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TTTGCAGATAATGTATAGTGTGTCACAAGGACATCGACCTGTTATTAATGAAGAAAGTTT 180

Qy   1024 gccatatgatatacctcaccgagcacgtat 1053
          ||||||||||||||||||||||||||||||
Db    181 GCCATATGATATACCTCACCGAGCACGTAT 210


RESULT   8
AF302127
LOCUS       AF302127       3559 bp    mRNA            ROD       08-NOV-2000
DEFINITION  Mus musculus PKC-regulated kinase PKK mRNA, complete cds.
ACCESSION   AF302127
VERSION     AF302127.1  GI:11120430
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3559)
  AUTHORS   Chen,L., Haider,K., Cariappa,A., Rowitch,D. and Pillai,S.
```

```
  TITLE     PKK, a novel ankyrin repeat-containing protein kinase regulated by
            protein kinase C-beta
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3559)
  AUTHORS   Chen,L., Haider,K., Cariappa,A. and Pillai,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-SEP-2000) Cancer Center, Massachusetts General
            Hospital, 149 13th Street, Charlestown, MA 02129-2060, USA
FEATURES             Location/Qualifiers
     source          1. .3559
                     /organism="Mus musculus"
                     /strain="BALB/c"
                     /db_xref="taxon:10090"
     CDS             50. .2410
                     /codon_start=1
                     /product="PKC-regulated kinase PKK"
                     /protein_id="AAG30871.1"
                     /db_xref="GI:11120431"
                     /translation="MEGEGRGRWALGLLRTFDAGEFAGWEKVGSGGFGQVYKVRHVHW
                     KTWLAIKCSPSLHVDDRERMELLEEAKKMEMAKFRYILPVYRICQEPVGLVMEYMETG
                     SLEKLLASEPLPWDLRFRIVHETAVGMNFLHCMSPPLLHLDLKPANILLDAHYHVKIS
                     DFGLAKCNGMSHSQDLSMDGLFGTIAYLPPERIREKSRLFDTKHDVYSFAIVIWGVLT
                     QKKPFADEKNILHIMMKVVKGHRPELPPICRPRPRACASLIGLMQRCWHADPQVRPTF
                     QEITSETEDLCEKPDEEVKDLAHEPGEKSSLESKSEARPESSRLKRASAPPFDNDCSL
                     SELLSQLDSGISQTLEGPEELSRSSSECKLPSSSSGKRLSGVSSVDSAFSSRGSLSLS
                     FEREASTGDLGPTDIQKKKLVDAIISGDTSRLMKILQPQDVDLVLDSSASLLHLAVEA
                     GQEECVKWLLLNNANPNLTNRKGSTPLHMAVERKGRGIVELLLARKTSVNAKDEDQWT
                     ALHFAAQNGDEASTRLLLEKNASVNEVDFEGRTPMHVACQHGQENIVRTLLRRGVDVG
                     LQGKDAWLPLHYAAWQGHLPIVKLLAKQPGVSVNAQTLDGRTPLHLAAQRGHYRVARI
                     LIDLCSDVNICSLQAQTPLHVAAETGHTSTARLLLHRGAGKEALTSEGYTALHLAAQN
                     GHLATVKLLIEEKADVMARGPLNQTALHLAAARGHSEVVEELVSADLIDLSDEQGLSA
                     LHLAAQGRHSQTVETLLKHGAHINLQSLKFQGGQSSAATLLRRSKT"
BASE COUNT      826 a    901 c   1000 g    832 t
ORIGIN

  Query Match             4.9%;  Score 102.4;  DB 94;  Length 3559;
  Best Local Similarity   51.8%;  Pred. No. 1.4e-08;
  Matches  372;  Conservative    0;  Mismatches  316;  Indels   30;  Gaps    5;

Qy      287 ctgcccaccattccctaccacaaactcgccgacctgcgctacctgagccgcggcgcctct 346
            |||| |||| |   |  |  | || |||| | |  |    |  || ||   |||| ||
Db       89 CTGCGCACCTTCGACGCCGGCGAATTCGCAGGCTGGGAGAAGGTGGGCTCGGGCGGCTTC 148

Qy      347 ggcactgtgtcgtccgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctg 406
            ||    ||||     |  ||||| |   |||||        | | ||  | |||  |  |
Db      149 GGCAGGTGTACAAGGTGCGCCATGTGCACTGGAAGACGTGGCTCGCGATCAAGTGCTCG 208

Qy      407 cacatccacactccgctgctcgacagtgaaagaaaggatgttttaagagaagctgaaatt 466
            | ||  |     |   |   |||||| ||| ||| |||  |  |    |||||| | |
Db      209 CCCAGTC---TGCACGTCGACGACAGGGAACGAATGGAGCTCCTGGAGGAAGCTAAGAAG 265

Qy      467 ttacacaaagctagatttagttacattttccaattttgggaatttgcaatgagcctgaa 526
             |  | |  || |  ||  | |||||| | ||  | |   | || ||| | || ||||
Db      266 ATGGAGATGGCCAAGTTCCGATACATTCTACCTGTGTACCGCATATGCCAGGAACCTG-- 323

Qy      527 tttttgggaatagttactgaatacatgccaaatggatcattaaatgaactcctacatagg 586
                | ||  | || |  || ||||||   |  || ||  |  |  | || ||
Db      324 ----TCGGCTTGGTCATGGAGTACATGGAGACAGGCTCCCTGGAAAAGCTGCTGGCCTCA 379

Qy      587 aaaactgaatatcctgatgttgcttggccattgagatttcgcatcctgcatgaaattgcc 646
             |  |              |  |||||    || | ||||||||| |||| || |  |||
Db      380 GAGCCAT------------TGCCTTGGGACCTGCGCTTTCGCATCGTGCACGAGACAGCC 427

Qy      647 cttggtgtaaattacctgcacaatatgactcctcctttacttcatcatgacttgaagact 706
             | ||  | || | ||||||    ||| |||| ||  | || || |  ||| ||||| |
Db      428 GTGGGCATGAACTTCCTGCATTGCATGTCTCCGCCACTGCTGCACCTAGACCTGAAGCCA 487

Qy      707 cagaatatcttattggacaatgaatttcatgttaagattgcagattttggtttatcaaag 766
              ||| ||| |  ||||     | |  ||||| |||||| | || |||||  |  | |||
Db      488 GCGAACATCCTGCTGGATGCCCACTACCATGTCAAGATTTCTGACTTTGGGCTGGCCAAG 547

Qy      767 tggcgcatgatgtccctctcacagtcacgaagtagcaaatctgcaccagaaggagggaca 826
```

```
            ||       ||||||| ||| |||   |  || |   |                || |||
Db      548 TGCAATGGCATGTCCCACTCTCAGGACCTCAGCATGGATGGCCTGTTT------GGTACA 601

Qy      827 attatttatatgccacctgaaaactatgaacctggacaaaaatcaagggccagtatcaag 886
            ||   |||  | || || ||      |   | ||   | |       |      | |||
Db      602 ATCGCTTACCTCCCTCCAGAGCGAATT---CGTGAGAAGAGCCGCTTGTTTGACACCAAA 658

Qy      887 cacgatatatatagctatgcagttatcacatgggaagtgttatccagaaaacagcctttt 946
            || ||| |||| ||||  ||  || | |  ||||  ||| |  |    ||  |||| |||
Db      659 CATGATGTATACAGCTTCGCCATTGTGATCTGGGGTGTGCTTACACAGAAGAAGCCATTT 718

Qy      947 gaagatgtcaccaatcctttgcagataatgtatagtgtgtcacaaggacatcgacctg 1004
            | |||||  |  ||     | || || |||   |  |||  | | || || || || |
Db      719 GCAGATGAAAAGAACATCCTACACATCATGATGAAAGTGGTAAAGGGCCACCGCCCAG 776


RESULT   9
HSA278016
LOCUS       HSA278016    3879 bp    mRNA            PRI       19-NOV-2000
DEFINITION  Homo sapiens mRNA for protein kinase (dik gene).
ACCESSION   AJ278016
VERSION     AJ278016.1  GI:9886710
KEYWORDS    dik gene; protein kinase.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3879)
  AUTHORS   Bahr,C., Rohwer,A., Stempka,L., Rincke,G., Marks,F. and
            Gschwendt,M.
  TITLE     DIK, a Novel Protein Kinase That Interacts with Protein Kinase
            Cdelta. CLONING, CHARACTERIZATION, AND GENE ANALYSIS
  JOURNAL   J. Biol. Chem. 275 (46), 36350-36357 (2000)
   PUBMED   10948194
REFERENCE   2  (bases 1 to 3879)
  AUTHORS   Gschwendt,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-MAY-2000) Gschwendt M., Biochemistry of
            Tissue-specific Regulation, German Cancer Research Center, Im
            Neuenheimer Feld 280, D-69151 Heidelberg, GERMANY
FEATURES             Location/Qualifiers
     source          1. .3879
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HaCaT"
     gene            49. .3843
                     /gene="dik"
     CDS             49. .2403
                     /gene="dik"
                     /function="protein phosphorylation"
                     /codon_start=1
                     /evidence=experimental
                     /product="protein kinase"
                     /protein_id="CAC04247.1"
                     /db_xref="GI:9886711"
                     /translation="MEGDGGTPWALALLRTFDAGEFTGWEKVGSGGFGQVYKVRHVHW
                     KTWLAIKCSPSLHVDDRERMELLEEAKKMEMAKFRYILPVYGICREPVGLVMEYMETG
                     SLEKLLASEPLPWDLRFRIIHETAVGMNFLHCMAPPLLHLDLKPANILLDAHYHVKIS
                     DFGLAKCNGLSHSHDLSMDGLFGTIAYLPPERIREKSRLFDTKHDVYSFAIVIWGVLT
                     QKKPFADEKNILHIMVKVVKGHRPELPPVCRARPRACSHLIRLMQRCWQGDPRVRPTF
                     QEITSETEDLCEKPDDEVKETAHDLDVKSPPEPRSEVVPARLKRASAPTFDNDYSLSE
                     LLSQLDSGVSQAVEGPEELSRSSSESKLPSSGSGKRLSGVSSVDSAFSSRGSLSLSFE
                     REPSTSDLGTTDVQKKKLVDAIVSGDTSKLMKILQPQDVDLALDSGASLLHLAVEAGQ
                     EECAKWLLLNNANPNLSNRRGSTPLHMAVERRVRGVVELLLARKISVNAKDEDQWTAL
                     HFAAQNGDESSTRLLLEKNASVNEVDFEGRTPMHVACQHGQENIVRILLRRGVDVSLQ
                     GKDAWLPLHYAAWQGHLPIVKLLAKQPGVSVNAQTLDGRTPLHLAAQRGHYRVARILI
                     DLCSDVNVCSLLAQTPLHVAAETGHTSTARLLLHRGAGKKAVTSDGYTALHLAARNGH
                     LATVKLLVEEKADVLARGPLNQTALHLAAAHGHSEVVEELVSADVIDLFDEQGLSALH
                     LAAQGRHAQTVETLLRHGAHINLQSLKFQGGHGPAATLLRRSKT"
     polyA_signal    3838. .3843
                     /gene="dik"
BASE COUNT      807 a   1074 c   1178 g    820 t
ORIGIN
```

```
  Query Match             4.7%;  Score 97.6;  DB 92;  Length 3879;
  Best Local Similarity   51.4%;  Pred. No. 1e-07;
  Matches  369;  Conservative    0;  Mismatches  319;  Indels   30;  Gaps    5;

Qy     287 ctgcccaccattccctaccacaaactcgccgacctgcgctacctgagccgcggcgcctct 346
           |||| |||| |   |     | |  || | | |  |    |  || ||   |||| ||
Db      88 CTGCGCACCTTCGACGCGGGCGAGTTCACGGGCTGGGAGAAGGTGGGCTCGGGCGGCTTC 147

Qy     347 ggcactgtgtcgtccgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctg 406
           ||    ||||     |  ||||| |   |||||     |  | ||||| | |||  |  |
Db     148 GGGCAGGTGTACAAGGTGCGCCATGTCCACTGGAAGACCTGGCTGGCCATCAAGTGCTCG 207

Qy     407 cacatccacactccgctgctcgacagtgaaagaaaggatgttttaagagaagctgaaatt 466
           | || ||     |   |   |||||| ||  | | |||  ||||   ||||||  | |
Db     208 CCCAGCC---TGCACGTCGACGACAGGGAGCGCATGGAGCTTTTGGAAGAAGCCAAGAAG 264

Qy     467 ttacacaaagctagatttagttacatttttccaattttgggaatttgcaatgagcctgaa 526
            |  | |  || |  ||| | |||||  | ||  | |  || || |||    || ||||
Db     265 ATGGAGATGGCCAAGTTTCGCTACATCCTGCCTGTGTATGGCATCTGCCGCGAACCTG-- 322

Qy     527 tttttgggaatagttactgaatacatgccaaatggatcattaaatgaactcctacatagg 586
               | ||  | || |  || ||||||   |  || ||  |  |  | || ||   |  |
Db     323 ----TCGGCCTGGTCATGGAGTACATGGAGACGGGCTCCCTGGAAAAGCTGCTGGCTTCG 378

Qy     587 aaaactgaatatcctgatgttgcttggccattgagatttcgcatcctgcatgaaattgcc 646
            |  |              |  | |||    |  | || || ||| | || || |  ||
Db     379 GAGCCAT------------TGCCATGGGATCTCCGGTTCCGAATCATCCACGAGACGGCG 426

Qy     647 cttggtgtaaattacctgcacaatatgactcctcctttacttcatcatgacttgaagact 706
            | ||  | || | |||||||   ||| | || ||  | || || |  ||| | ||| |
Db     427 GTGGGCATGAACTTCCTGCACTGCATGGCCCCGCCACTCCTGCACCTGGACCTCAAGCCC 486

Qy     707 cagaatatcttattggacaatgaatttcatgttaagattgcagattttggtttatcaaag 766
             ||| ||| |  ||||     | |  || || |||||| | ||||||||| |  | |||
Db     487 GCGAACATCCTGCTGGATGCCCACTACCACGTCAAGATTTCTGATTTTGGTCTGGCCAAG 546

Qy     767 tggcgcatgatgtccctctcacagtcacgaagtagcaaatctgcaccagaaggagggaca 826
           ||   |  | |||||| ||| ||    |  || |   |                  || |||
Db     547 TGCAACGGGCTGTCCCACTCGCATGACCTCAGCATGGATGGCCTGTTT------GGCACA 600

Qy     827 attatttatatgccacctgaaaactatgaacctggacaaaaatcaagggccagtatcaag 886
           ||    ||  | || || ||   |         |   | |          |   | ||||
Db     601 ATCGCCTACCTCCCTCCAGAGCGCATCAGG---GAGAAGAGCCGGCTCTTCGACACCAAG 657

Qy     887 cacgatatatatagctatgcagttatcacatgggaagtgttatccagaaaacagcctttt 946
           |||||| |||| |||| |||  |  |||  ||||  ||| |  |    ||  |||| |||
Db     658 CACGATGTATACAGCTTTGCGATCGTCATCTGGGGCGTGCTCACACAGAAGAAGCCGTTT 717

Qy     947 gaagatgtcaccaatcctttgcagataatgtatagtgtgtcacaaggacatcgacctg 1004
           | |||||  |  ||     |||| || |||   |  |||    | || || || || |
Db     718 GCAGATGAGAAGAACATCCTGCACATCATGGTGAAGGTGGTGAAGGGCCACCGCCCCG 775


RESULT  10
I14046
LOCUS       I14046       3033 bp    DNA             PAT       26-SEP-1995
DEFINITION  Sequence 1 from patent US 5444166.
ACCESSION   I14046
VERSION     I14046.1  GI:996469
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3033)
  AUTHORS   Ecker,J.R. and Kieber,J.J.
  TITLE     Constitutive triple response gene and mutations
  JOURNAL   Patent: US 5444166-A 1 22-AUG-1995;
FEATURES             Location/Qualifiers
     source          1. .3033
                     /organism="unknown"
BASE COUNT      797 a    605 c    751 g    880 t
```

```
ORIGIN

  Query Match             3.9%;  Score 81.4;  DB 10;  Length 3033;
  Best Local Similarity   50.1%;  Pred. No. 7.6e-05;
  Matches  202;  Conservative    0;  Mismatches  201;  Indels    0;  Gaps    0;

Qy     360 ccgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctgcacatccacactc 419
           | | || ||  || || ||||  | |  ||  |  |    | |  | | |||  |
Db    1808 CTGTCCACCGTGCTGAGTGGCATGGCTCGGATGTTGCTGTGAAAATTCTCATGGAGCAAG 1867

Qy     420 cgctgctcgacagtgaaagaaaggatgttttaagagaagctgaaattttacacaaagcta 479
              | |  |           || ||  | |||||||| | ||  ||  |  |      |
Db    1868 ACTTCCATGCTGAGCGTGTTAATGAGTTCTTAAGAGAGGTTGCGATAATGAAACGCCTTC 1927

Qy     480 gatttagttacatttttccaattttgggaatttgcaatgagcctgaatttttgggaatag 539
           |      | ||||| |||    |  |||        || | | |||  |  ||||  |||||
Db    1928 GCCACCCTAACATTGTTCTCTTCATGGGTGCGGTCACTCAACCTCCAAATTTGTCAATAG 1987

Qy     540 ttactgaatacatgccaaatggatcattaaatgaactcctacataggaaaactgaatatc 599
           | || |||||  || |||  ||    ||| |   |||  | ||||  |     | |
Db    1988 TGACAGAATATTTGTCAAGAGGTAGTTTATACAGACTTTTGCATAAAAGTGGAGCAAGGG 2047

Qy     600 ctgatgttgcttggccattgagatttcgcatcctgcatgaaattgcccttggtgtaaatt 659
              |  | | |  |  |    |  |  | ||     ||||  | ||    ||  | ||||
Db    2048 AGCAATTAGATGAGAGACGTCGCCTGAGTATGGCTTATGATGTGGCTAAGGGAATGAATT 2107

Qy     660 acctgcacaatatgactcctcctttacttcatcatgacttgaagactcagaatatcttat 719
           | || ||||||   | ||||||  |  | |||   ||  | ||  |||  ||  | ||
Db    2108 ATCTTCACAATCGCAATCCTCCAATTGTGCATAGAGATCTAAAATCTCCAAACTTATTGG 2167

Qy     720 tggacaatgaatttcatgttaagattgcagattttggtttatc 762
           | |||||  ||| |   || ||| ||   ||||||||| | ||
Db    2168 TTGACAAAAAATATACAGTCAAGGTTTGTGATTTTGGTCTCTC 2210


RESULT  11
I35764
LOCUS       I35764       3033 bp    DNA             PAT       13-MAY-1997
DEFINITION  Sequence 1 from patent US 5602322.
ACCESSION   I35764
VERSION     I35764.1  GI:2087615
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3033)
  AUTHORS   Ecker,J.R. and Kieber,J.J.
  TITLE     Constitututive triple response gene and mutations
  JOURNAL   Patent: US 5602322-A 1 11-FEB-1997;
FEATURES             Location/Qualifiers
     source          1. .3033
                     /organism="unknown"
BASE COUNT      797 a    605 c    751 g    880 t
ORIGIN


  Query Match             3.9%;  Score 81.4;  DB 10;  Length 3033;
  Best Local Similarity   50.1%;  Pred. No. 7.6e-05;
  Matches  202;  Conservative    0;  Mismatches  201;  Indels    0;  Gaps    0;

Qy     360 ccgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctgcacatccacactc 419
           | | || ||  || || ||||  | |  ||  |  |    | |  | | |||  |
Db    1808 CTGTCCACCGTGCTGAGTGGCATGGCTCGGATGTTGCTGTGAAAATTCTCATGGAGCAAG 1867

Qy     420 cgctgctcgacagtgaaagaaaggatgttttaagagaagctgaaattttacacaaagcta 479
              | |  |           || ||  | |||||||| | ||  ||  |  |      |
Db    1868 ACTTCCATGCTGAGCGTGTTAATGAGTTCTTAAGAGAGGTTGCGATAATGAAACGCCTTC 1927

Qy     480 gatttagttacatttttccaattttgggaatttgcaatgagcctgaatttttgggaatag 539
           |      | ||||| |||    |  |||        || | | |||  |  ||||  |||||
Db    1928 GCCACCCTAACATTGTTCTCTTCATGGGTGCGGTCACTCAACCTCCAAATTTGTCAATAG 1987
```

```
Qy     540 ttactgaatacatgccaaatggatcattaaatgaactcctacataggaaaactgaatatc 599
           | || |||||  || |||  ||    ||| |   |||  | ||||  |     | |
Db    1988 TGACAGAATATTTGTCAAGAGGTAGTTTATACAGACTTTTGCATAAAAGTGGAGCAAGGG 2047

Qy     600 ctgatgttgcttggccattgagatttcgcatcctgcatgaaattgcccttggtgtaaatt 659
              |  | | |  |  |    |  |  | ||     ||||  | ||    ||  | ||||
Db    2048 AGCAATTAGATGAGAGACGTCGCCTGAGTATGGCTTATGATGTGGCTAAGGGAATGAATT 2107

Qy     660 acctgcacaatatgactcctcctttacttcatcatgacttgaagactcagaatatcttat 719
           | || ||||||   | ||||||  |  | |||   ||  | ||  |||  ||  | ||
Db    2108 ATCTTCACAATCGCAATCCTCCAATTGTGCATAGAGATCTAAAATCTCCAAACTTATTGG 2167

Qy     720 tggacaatgaatttcatgttaagattgcagattttggtttatc 762
           | |||||  ||| |   || ||| ||   ||||||||| | ||
Db    2168 TTGACAAAAAATATACAGTCAAGGTTTGTGATTTTGGTCTCTC 2210


RESULT  12
ATHCTR1A
LOCUS       ATHCTR1A     3033 bp    mRNA            PLN       27-APR-1993
DEFINITION  Arabidopsis thaliana negative regulator of the ethylene response
            pathway (CTR1) mRNA, complete cds.
ACCESSION   L08789
VERSION     L08789.1  GI:166679
KEYWORDS    ethylene response pathway; negative regulator; protein kinase.
SOURCE      Arabidopsis thaliana (strain Columbia) dark grown seedling (3 days
            old) hypocotyl and cotyledon cDNA to mRNA.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
            Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 3033)
  AUTHORS   Kieber,J.J., Rothenberg,M., Roman,G., Feldmann,K.A. and Ecker,J.R.
  TITLE     CTR1, a negative regulator of the ethylene response pathway in
            Arabidopsis, encodes a member of the Raf family of protein kinases
  JOURNAL   Cell 72, 427-441 (1993)
  MEDLINE   93161417
FEATURES             Location/Qualifiers
     source          1. .3033
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /dev_stage="dark grown seedling (3 days old)"
                     /tissue_type="hypocotyl and cotyledon"
     gene            118. .2583
                     /gene="CTR1"
     CDS             118. .2583
                     /gene="CTR1"
                     /codon_start=1
                     /product="protein kinase"
                     /protein_id="AAA32779.1"
                     /db_xref="GI:166680"
                     /translation="MEMPGRRSNYTLLSQFSDDQVSVSVTGAPPPHYDSLSSENRSNH
                     NSGNTGKAKAERGGFDWDPSGGGGGDHRLNNQPNRVGNNMYASSLGLQRQSSGSSFGE
                     SSLSGDYYMPTLSAAANEIESVGFPQDDGFRLGFGGGGGDLRIQMAADSAGGSSSGKS
                     WAQQTEESYQLQLALALRLSSEATCADDPNFLDPVPDESALRTSPSSAETVSHRFWVN
                     GCLSYYDKVPDGFYMMNGLDPYIWTLCIDLHESGRIPSIESLRAVDSGVDSSLEAIIV
                     DRRSDPAFKELHNRVHDISCSCITTKEVVDQLAKLICNRMGGPVIMGEDELVPMWKEC
                     IDGLKEIFKVVVPIGSLSVGLCRHRALLFKVLADIIDLPCRIAKGCKYCNRDDAASCL
                     VRFGLDREYLVDLVGKPGHLWEPDSLLNGPSSISISSPLRFPRPKPVEPAVDFRLLAK
                     QYFSDSQSLNLVFDPASDDMGFSMFHRQYDNPGGENDALAENGGGSLPPSANMPPQNM
                     MRASNQIEAAPMNAPPISQPVPNRANRELGLDGDDMDIPWCDLNIKEKIGAGSFGTVH
                     RAEWHGSDVAVKILMEQDFHAERVNEFLREVAIMKRLRHPNIVLFMGAVTQPPNLSIV
                     TEYLSRGSLYRLLHKSGAREQLDERRRLSMAYDVAKGMNYLHNRNPPIVHRDLKSPNL
                     LVDKKYTVKVCDFGLSRLKASTFLSSKSAAGTPEWMAPEVLRDEPSNEKSDVYSFGVI
                     LWELATLQQPWGNLNPAQVVAAVGFKCKRLEIPRNLNPQVAAIIEGCWTNEPWKRPSF
                     ATIMDLLRPLIKSAVPPPNRSDL"
BASE COUNT      797 a    605 c    751 g    880 t
ORIGIN


  Query Match             3.9%;  Score 81.4;  DB 13;  Length 3033;
  Best Local Similarity   50.1%;  Pred. No. 7.6e-05;
  Matches  202;  Conservative    0;  Mismatches  201;  Indels    0;  Gaps    0;

Qy     360 ccgcccgccacgcagactggcgcgtccaggtggccgtgaagcacctgcacatccacactc 419
           | | || ||  || || ||||  | |  ||  |  |    | |  | | |||  |
Db    1808 CTGTCCACCGTGCTGAGTGGCATGGCTCGGATGTTGCTGTGAAAATTCTCATGGAGCAAG 1867

Qy     420 cgctgctcgacagtgaaagaaaggatgttttaagagaagctgaaattttacacaaagcta 479
              | |  |           || ||  | |||||||| | ||  ||  |  |      |
Db    1868 ACTTCCATGCTGAGCGTGTTAATGAGTTCTTAAGAGAGGTTGCGATAATGAAACGCCTTC 1927

Qy     480 gatttagttacattttccaatttgggaatttgcaatgagcctgaatttttgggaatag 539
           |      | ||||| |||   |  ||||      || | | |||  |  ||||  |||||
Db    1928 GCCACCCTAACATTGTTCTCTTCATGGGTGCGGTCACTCAACCTCCAAATTTGTCAATAG 1987

Qy     540 ttactgaatacatgccaaatggatcattaaatgaactcctacataggaaaactgaatatc 599
           | || |||||  || |||  ||    ||| |   |||  | ||||  |     | |
Db    1988 TGACAGAATATTTGTCAAGAGGTAGTTTATACAGACTTTTGCATAAAAGTGGAGCAAGGG 2047

Qy     600 ctgatgttgcttggccattgagatttcgcatcctgcatgaaattgcccttggtgtaaatt 659
              |  | | |  |  |    |  |  | ||     ||||  | ||    ||  | ||||
Db    2048 AGCAATTAGATGAGAGACGTCGCCTGAGTATGGCTTATGATGTGGCTAAGGGAATGAATT 2107

Qy     660 acctgcacaatatgactcctcctttacttcatcatgacttgaagactcagaatatcttat 719
           | || ||||||   | ||||||  |  | |||   ||  | ||  |||  ||  | ||
Db    2108 ATCTTCACAATCGCAATCCTCCAATTGTGCATAGAGATCTAAAATCTCCAAACTTATTGG 2167

Qy     720 tggacaatgaatttcatgttaagattgcagattttggtttatc 762
           | |||||  ||| |   || ||| ||   ||||||||| | ||
Db    2168 TTGACAAAAAATATACAGTCAAGGTTTGTGATTTTGGTCTCTC 2210


RESULT  13
AF110519
LOCUS       AF110519     2767 bp    mRNA            PLN       28-JAN-1999
DEFINITION  Lycopersicon esculentum ethylene-inducible CTR1-like protein kinase
            mRNA, complete cds.
ACCESSION   AF110519
VERSION     AF110519.1  GI:4193949
KEYWORDS    .
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
            I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 2767)
  AUTHORS   Kannan,P. and Giovannoni,J.J.
  TITLE     Ethylene-inducible tomato CTR1-like protein kinase
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2767)
  AUTHORS   Kannan,P. and Giovannoni,J.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-DEC-1998) Horticultural Sciences, Texas A&M
            University, Mail Stop 2133, College Station, TX 77843-2133, USA
FEATURES             Location/Qualifiers
     source          1. .2767
                     /organism="Lycopersicon esculentum"
                     /cultivar="Ailsa Craig"
                     /db_xref="taxon:4081"
                     /chromosome="2"
                     /map="between TG33 and CT106A"
     CDS             67. .2487
                     /note="protein kinase homolog; ethylene and fruit ripening
                     inducible CTR1-like protein kinase; TCTR1v"
                     /codon_start=1
                     /product="ethylene-inducible CTR1-like protein kinase"
                     /protein_id="AAD10057.1"
                     /db_xref="GI:4193950"
                     /translation="MSGRRSSYTLLNQIPNDNFFQPPAPKFSAGAGVLPYGESSSAEK
                     NRGKVFDLDLMDQRMMQSHNRVGSFRVPGSIGSQKQSTEGSFRGSSLSGENYVGTSFG
                     HKNEGCGSSVARNWAQQTEESYQLQLALAIRLSSEATCADSPNFLDPVTDVLASRDSD
                     STASAVTMSHRLWINGCMSYFDKVPDGFYWIYGMDPYVWALCSVVQESGRIPSIESLK
                     AVDPSKAPSVEVILIVRCNDLSLKELQNRIHSISPSCITTKEAVDQLAKLVCDHMGGA
                     APAGEEELVSMSKGCSNDLKDRFGTIVLPIGSLSVGLCRHRALLFKVLADIIDLPCRI
```

```
                     AKGCKYCNSSDASSCLVRFEHDREYLVDLIGKPGVLSEPDSLVNGPFSISIPSPLRFP
                     RYRQVEPTTDFRAAADGDAGQSDRSCIDRNNVVSSSSNRDEISQLPPPPLNAWKKGRD
                     KESQLSKMYNPRSMLNPVNMDEDQVLVKHVPPFREDAQSPMTRPDTVNDTRFLAGGGH
                     VVSAIPSEELDLDVEEFNIPWNDLVLMEKIGAGSFGTVHRGDWHGSDVAVKILMEQDF
                     HAERLKEFLREVAIMKRLRHPNIVLFMGAVIQPPNLSIVTEYLSRGSLYRLLHKPGAK
                     KVLDERRPLCLAYDVANGMNYLHKRNPPIVHRDLKSPNLLVDKKYTVKICDFGLSRFK
                     ANTFLSSKTAAGTPEWMAPEVIRDEPSNEKSDVYSFGVILWELATLQQPWNKLNPPQV
                     IAAVGFNRKKLDIPSVLNPRVAIIIEACWANEPWKRPSFSTIMDMLRPHLKSPLPPPG
                     HTDMQLLS"
BASE COUNT      728 a    554 c    659 g    826 t
ORIGIN


  Query Match             3.7%;  Score 78.2;  DB 12;  Length 2767;
  Best Local Similarity   52.6%;  Pred. No. 0.00028;
  Matches  170;  Conservative    0;  Mismatches  153;  Indels    0;  Gaps    0;

Qy    440 aaggatgttttaagagaagctgaaattttacacaaagctagatttagttacatttttcca 499
          |||||  |||| || || | || |||| |  |        ||  |    | ||| | |
Db   1780 AAGGAATTTTTGAGGGAGGTTGCAATTATGAAGCGGTTGCGACATCCAAATATTGTACTT 1839

Qy    500 attttgggaatttgcaatgagcctgaatttttgggaatagttactgaatacatgccaaat 559
           || ||||   |  || | ||||   |  ||||   ||||| || |||||  |  | |
Db   1840 TTTATGGGTGCTGTCATTCAGCCACCAAATTTGTCCATAGTCACGGAATATTTATCGAGA 1899

Qy    560 ggatcattaaatgaactcctacataggaaaactgaatatcctgatgttgcttggccattg 619
          ||    ||| ||  ||| || ||||       ||   |    |   | | |
Db   1900 GGTAGCTTATATAGACTTCTTCATAAACCTGGTGCGAAAAAAGTGTTGGATGAAAGGCGT 1959

Qy    620 agatttcgcatcctgcatgaaattgcccttggtgtaaattacctgcacaatatgactcct 679
              |  |  |     | ||  | |||  |||  | ||||| || |||||    | ||||
Db   1960 CCCCTGTGTTTGGCTTACGATGTGGCCAATGGGATGAATTATCTTCACAAACGCAATCCT 2019

Qy    680 cctttacttcatcatgacttgaagactcagaatatcttattggacaatgaatttcatgtt 739
          ||  |  | || |  || || ||  |||  ||| |  || | |||||  ||| |   ||
Db   2020 CCCATTGTGCACCGAGATTTAAAATCTCCAAATCTTCTAGTAGACAAAAAATATACAGTG 2079

Qy    740 aagattgcagattttggtttatc 762
          |||||    ||||||||| | ||
Db   2080 AAGATCTGTGATTTTGGTCTTTC 2102


RESULT  14
AF096250
LOCUS       AF096250     2829 bp    mRNA            PLN       02-AUG-1999
DEFINITION  Lycopersicon esculentum ethylene-responsive protein kinase TCTR1
            (ER50) mRNA, complete cds.
ACCESSION   AF096250
VERSION     AF096250.1  GI:5669641
KEYWORDS    .
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
            I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 2829)
  AUTHORS   Zegzouti,H., Jones,B., Frasse,P., Marty,C., Maitre,B., Latch,A.,
            Pech,J.C. and Bouzayen,M.
  TITLE     Ethylene-regulated gene expression in tomato fruit:
            characterization of novel ethylene-responsive and ripening-related
            genes isolated by differential display
  JOURNAL   Plant J. 18 (6), 589-600 (1999)
  MEDLINE   99348164
REFERENCE   2  (bases 1 to 2829)
  AUTHORS   Zegzouti,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-OCT-1998) Lab 'ETHYLENE', ENSAT-INRA, Avenue de
            l'Agrobiopole BP 107, Auzeville Tolosan, Castanet Tolosan cedex
            31326, France
FEATURES             Location/Qualifiers
     source          1. .2829
                     /organism="Lycopersicon esculentum"
                     /cultivar="Evita"
                     /db_xref="taxon:4081"
                     /tissue_type="green fruit"
                     /dev_stage="late immature"
                     /note="ethylene-treated"
     gene            1. .2829
                     /gene="ER50"
                     /note="ethylene-inducible"
     5'UTR           1. .52
                     /gene="ER50"
     CDS             53. .2542
                     /gene="ER50"
                     /note="serine/threonine kinase; similar to Arabidopsis
                     thaliana negative regulator of the ethylene response
                     pathway encoded by GenBank Accession Number L08789"
                     /codon_start=1
                     /product="ethylene-responsive protein kinase TCTR1"
                     /protein_id="AAD46406.1"
                     /db_xref="GI:5669642"
                     /translation="MSGRRSSYTLLNQIPNDNFFQPPAPKFSAGAGVVPYGESSSAEK
                     NRGKVFDLDLMDQRMMQSHNRVGSFRVPGSIGSQRQSSEGSFGGSSLSGENYVGTSFG
                     HKNEGCGSSVARSWAQQTEESYQLQLALAIRLSSEATCADSPNFLDPVTDVLASRDSD
                     STASAVTMSHRLWINGCMSYFDKVPDGFYWIYGMDPYVWALCSVVQESGRIPSIESLR
                     AVDPSKAPSVEVILIDRCNDLSLKELQNRIHSISPSCITTKEAVDQLAKLVCDHMGGA
                     APAGEEELVSMSKGCSNDLKDRFGTIVLPIGSLSVGLCRHRALLFKVLADIIDLPCRI
                     AKGCKYCNSSDASSCLVRFEHDREYLVDLIGKPGVLSEPDSLLNGPSSISIPSPLRFP
                     RYRQVEPTTDFRSLAKQYFLDSQSLNLLFDDSSAGAAADGDAGQSDRSCIDRNNVVSS
                     SSNRDEISQLPLPPLNAWKKGRDKESQLSKMYNPRSMLNPVNMDEDQVLVKHVPPFRE
                     DAQSPMTRPDTVNDTRFLAGGGHVVSAIPSEELDLDVEEFNIPWNDLILMEKIGAGSF
                     GTVHRGDWHGSDVAVKILMEQDFHAERLKEFLREVAIMKRLRHPNIVLFMGAVIQPPN
                     LSIVTEYLSRGSLYRLLHKPGAREVLDERRRLCMAYDVANGMNYLHKRNPPIVHRDLK
                     SPNLLVDKKYTVKICDFGLSRFKANTFLSSKTAAGTPEWMAPEVIRDEPSNEKSDVYS
                     FGVILWELATLQQPWNKLNPPQVIAAVGFNRKRLDIPSDLNPQVAIIIEACWANEPWK
                     RPSFSTIMDMLRPHLKSPLPPPGHTDMQLLS"
     3'UTR           2543. .2829
                     /gene="ER50"
BASE COUNT      739 a    561 c    684 g    845 t
ORIGIN


  Query Match             3.7%;  Score 78.2;  DB 12;  Length 2829;
  Best Local Similarity   52.6%;  Pred. No. 0.00028;
  Matches  170;  Conservative    0;  Mismatches  153;  Indels    0;  Gaps    0;

Qy    440 aaggatgttttaagagaagctgaaattttacacaaagctagatttagttacatttttcca 499
          |||||  |||| || || | || |||| |  |        ||  |    | ||| | |
Db   1835 AAGGAATTTTTGAGGGAGGTTGCAATTATGAAGCGGTTGCGACATCCAAATATTGTACTT 1894

Qy    500 attttgggaatttgcaatgagcctgaatttttgggaatagttactgaatacatgccaaat 559
           || ||||   |  || | ||||   |  ||||   ||||| || |||||  |  | |
Db   1895 TTTATGGGTGCTGTCATTCAGCCACCAAATTTGTCCATAGTCACGGAATATTTATCGAGA 1954

Qy    560 ggatcattaaatgaactcctacataggaaaactgaatatcctgatgttgcttggccattg 619
          ||    ||| ||  ||| || ||||       ||        |   | | |
Db   1955 GGTAGCTTATATAGACTTCTTCATAAACCTGGTGCGAGAGAGGTGTTGGATGAAAGGCGT 2014

Qy    620 agatttcgcatcctgcatgaaattgcccttggtgtaaattacctgcacaatatgactcct 679
           |  |  | ||     | ||  | ||   |||  | ||||| || |||||    | ||||
Db   2015 CGCCTGTGTATGGCTTACGATGTGGCAAATGGGATGAATTATCTTCACAAACGCAATCCT 2074

Qy    680 cctttacttcatcatgacttgaagactcagaatatcttattggacaatgaatttcatgtt 739
          ||  |  | || |  || || ||  |||  ||| |  || | |||||  ||| |   ||
Db   2075 CCCATTGTGCACCGAGATTTAAAATCTCCAAATCTTCTAGTAGACAAAAAATATACAGTG 2134

Qy    740 aagattgcagattttggtttatc 762
          |||||    ||||||||| | ||
Db   2135 AAGATCTGTGATTTTGGTCTTTC 2157


RESULT  15
AF110518
LOCUS       AF110518     2836 bp    mRNA            PLN       28-JAN-1999
DEFINITION  Lycopersicon esculentum ethylene-inducible CTR1-like protein kinase
            mRNA, complete cds.
```

```
ACCESSION   AF110518
VERSION     AF110518.1  GI:4193947
KEYWORDS    .
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
            I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 2836)
  AUTHORS   Kannan,P. and Giovannoni,J.J.
  TITLE     Ethylene-inducible tomato CTR1-like protein kinase
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2836)
  AUTHORS   Kannan,P. and Giovannoni,J.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-DEC-1998) Horticultural Sciences, Texas A&M
            University, Mail Stop 2133, College Station, TX 77843-2133, USA
FEATURES             Location/Qualifiers
     source          1. .2836
                     /organism="Lycopersicon esculentum"
                     /cultivar="Ailsa Craig"
                     /db_xref="taxon:4081"
                     /chromosome="2"
                     /map="between TG33 and CT106A"
     CDS             67. .2556
                     /note="protein kinase homolog; ethylene and fruit ripening
                     inducible CTR1-like protein kinase; TCTR1"
                     /codon_start=1
                     /product="ethylene-inducible CTR1-like protein kinase"
                     /protein_id="AAD10056.1"
                     /db_xref="GI:4193948"
                     /translation="MSGRRSSYTLLNQIPNDNFFQPPAPKFSAGAGVLPYGESSSAEK
                     NRGKVFDLDLMDQRMMQSHNRVGSFRVPGSIGSQKQSTEGSFRGSSLSGENYVGTSFG
                     HKNEGCGSSVARNWAQQTEESYQLQLALAIRLSSEATCADSPNFLDPVTDVLASRDSD
                     STASAVTMSHRLWINGCMSYFDKVPDGFYWIYGMDPYVWALCSVVQESGRIPSIESLK
                     AVDPSKAPSVEVILIVRCNDLSLKELQNRIHSISPSCITTKEAVDQLAKLVCDHMGGA
                     APAGEEELVSMSKGCSNDLKDRFGTIVLPIGSLSVGLCRHRALLFKVLADIIDLPCRI
                     AKGCKYCNSSDASSCLVRFEHDREYLVDLIGKPGVLSEPDSLVNGPFSISIPSPLRFP
                     RYRQVEPTTDFRSLAKQYFLDSQSLNLLFDDSSAGAAADGDAGQSDRSCIDRNNVVSS
                     SSNRDEISQLPPPPLNAWKKGRDKESQLSKMYNPRSMLNPVNMDEDQVLVKHVPPFRE
                     DAQSPMTRPDTVNDTRFLAGGGHVVSAIPSEELDLDVEEFNIPWNDLVLMEKIGAGSF
                     GTVHRGDWHGSDVAVKILMEQDFHAERLKEFLREVAIMKRLRHPNIVLFMGAVIQPPN
                     LSIVTEYLSRGSLYRLLHKPGAKKVLDERRPLCLAYDVANGMNYLHKRNPPIVHRDLK
                     SPNLLVDKKYTVKICDFGLSRFKANTFLSSKTAAGTPEWMAPEVIRDEPSNEKSDVYS
                     FGVILWELATLQQPWNKLNPPQVIAAVGFNRKKLDIPSVLNPRVAIIIEACWANEPWK
                     RPSFSTIMDMLRPHLKSPLPPPGHTDMQLLS"
BASE COUNT      745 a    567 c    672 g    852 t
ORIGIN


  Query Match             3.7%;  Score 78.2;  DB 12;  Length 2836;
  Best Local Similarity   52.6%;  Pred. No. 0.00028;
  Matches  170;  Conservative    0;  Mismatches  153;  Indels    0;  Gaps    0;

Qy    440 aaggatgttttaagagaagctgaaattttacacaaagctagatttagttacatttttcca 499
          |||||  |||| || || | || |||| |  |        ||  |    | ||| | |
Db   1849 AAGGAATTTTTGAGGGAGGTTGCAATTATGAAGCGGTTGCGACATCCAAATATTGTACTT 1908

Qy    500 attttgggaatttgcaatgagcctgaatttttgggaatagttactgaatacatgccaaat 559
           || ||||   |  || | ||||   |  ||||   ||||| || |||||  |  | |
Db   1909 TTTATGGGTGCTGTCATTCAGCCACCAAATTTGTCCATAGTCACGGAATATTTATCGAGA 1968

Qy    560 ggatcattaaatgaactcctacataggaaaactgaatatcctgatgttgcttggccattg 619
          ||    ||| ||  ||| || ||||       ||   |    |   | | |
Db   1969 GGTAGCTTATATAGACTTCTTCATAAACCTGGTGCGAAAAAAGTGTTGGATGAAAGGCGT 2028

Qy    620 agatttcgcatcctgcatgaaattgcccttggtgtaaattacctgcacaatatgactcct 679
              |  |  |     | ||  | |||  |||  | ||||| || |||||    | ||||
Db   2029 CCCCTGTGTTTGGCTTACGATGTGGCCAATGGGATGAATTATCTTCACAAACGCAATCCT 2088

Qy    680 cctttacttcatcatgacttgaagactcagaatatcttattggacaatgaatttcatgtt 739
          ||  |  | || |  || || ||  |||  ||| |  || | |||||  ||| |   ||
Db   2089 CCCATTGTGCACCGAGATTTAAAATCTCCAAATCTTCTAGTAGACAAAAAATATACAGTG 2148
```

```
Qy    740 aagattgcagattttggtttatc 762
          |||||    ||||||||| | ||
Db   2149 AAGATCTGTGATTTTGGTCTTTC 2171


Search completed: June 13, 2001, 17:39:55
Job time: 9374 sec
```

```
                    GenCore version 4.5
             Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          June 14, 2001, 08:04:28 ; Search time 14.27 Seconds
                                          (without alignments)
                                          762.306 Million cell updates/sec

Title:           US-09-445-223-1
Perfect score:   2829
Sequence:        1 MNGEAICSALPTIPYHKLAD.........PEILVVSRSPSLNLLQNKSM 540

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters:      193259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
               1:  /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
               2:  /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
               3:  /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
               4:  /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
               5:  /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
               6:  /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
               %
Result         Query
  No.   Score  Match Length DB  ID                      Description
-----------------------------------------------------------------------
    1    2817   99.6   540  3  US-09-019-942-1          Sequence 1, Appli
    2     576   20.4   787  4  US-09-188-930-334        Sequence 334, App
    3     536   18.9   536  4  US-09-188-930-185        Sequence 185, App
    4     394   13.9   656  1  US-08-444-005-15         Sequence 15, Appl
    5     389   13.8   518  3  US-09-329-418-9          Sequence 9, Appli
    6     388   13.7   518  3  US-09-329-418-3          Sequence 3, Appli
    7   386.5   13.7   420  3  US-09-329-418-8          Sequence 8, Appli
    8     385   13.6   518  3  US-09-329-418-5          Sequence 5, Appli
    9     380   13.4   518  3  US-09-329-418-4          Sequence 4, Appli
   10     376   13.3   821  1  US-07-928-464-2          Sequence 2, Appli
   11     376   13.3   821  1  US-08-003-311B-2         Sequence 2, Appli
   12     376   13.3   821  1  US-08-261-432-2          Sequence 2, Appli
   13     376   13.3   821  5  PCT-US93-07347-2         Sequence 2, Appli
   14     372   13.1   261  3  US-09-329-418-6          Sequence 6, Appli
   15     369   13.0   263  3  US-09-035-706-5          Sequence 5, Appli
   16     369   13.0   263  3  US-08-955-841-5          Sequence 5, Appli
   17     369   13.0   671  4  US-09-132-118-2          Sequence 2, Appli
   18     369   13.0   709  1  US-08-444-005-17         Sequence 17, Appl
   19   292.5   10.3   668  1  US-08-205-018-2          Sequence 2, Appli
   20   292.5   10.3   859  1  US-08-395-580-2          Sequence 2, Appli
   21   292.5   10.3   859  5  PCT-US95-02792-2         Sequence 2, Appli
   22   291.5   10.3   455  3  US-09-221-235-5          Sequence 5, Appli
   23   291.5   10.3   455  4  US-09-221-928-5          Sequence 5, Appli
   24   291.5   10.3   455  4  US-09-221-527-5          Sequence 5, Appli
   25   291.5   10.3   455  4  US-09-221-236-5          Sequence 5, Appli
   26   291.5   10.3   455  4  US-09-221-416-5          Sequence 5, Appli
   27   291.5   10.3   455  4  US-09-221-245-5          Sequence 5, Appli
   28   291.5   10.3   455  4  US-09-163-115-5          Sequence 5, Appli
   29   291.5   10.3   455  4  US-09-221-528-5          Sequence 5, Appli
   30   291.5   10.3   455  4  US-09-593-553-5          Sequence 5, Appli
   31   291.5   10.3   455  4  US-09-221-237-5          Sequence 5, Appli
   32   289.5   10.2   999  2  US-08-473-553A-5         Sequence 5, Appli
   33   283.5   10.0   712  1  US-08-587-889-2          Sequence 2, Appli
   34   283.5   10.0   712  2  US-08-980-060-5          Sequence 5, Appli
   35   283.5   10.0   712  4  US-09-307-185-5          Sequence 5, Appli
   36   283.5   10.0   712  5  PCT-US96-09193-2         Sequence 2, Appli
   37   274.5    9.7   527  5  PCT-US95-05008-10        Sequence 10, Appl
   38     265    9.4   279  2  US-08-701-191A-37        Sequence 37, Appl
   39   264.5    9.3   271  3  US-09-035-706-6          Sequence 6, Appli
   40   264.5    9.3   271  3  US-08-955-841-6          Sequence 6, Appli
   41     264    9.3   326  1  US-08-571-758-11         Sequence 11, Appl
   42     264    9.3   326  1  US-08-909-984A-11        Sequence 11, Appl
   43     264    9.3   326  1  US-08-909-983-11         Sequence 11, Appl
   44   261.5    9.2   544  2  US-08-587-680A-25        Sequence 25, Appl
   45     260    9.2   588  1  US-08-391-615-5          Sequence 5, Appli

                              ALIGNMENTS

RESULT   1
US-09-019-942-1
; Sequence 1, Application US/09019942
; Patent No. 6033855
;  GENERAL INFORMATION:
;    APPLICANT:  Bertin, John
;    TITLE OF INVENTION:  GENES ENCODING CASPASE RECRUITMENT
;    TITLE OF INVENTION:  DOMAIN POLYPEPTIDES
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/019,942
;      FILING DATE:  06-FEB-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meiklejohn, Ph.D., Anita L.
;      REGISTRATION NUMBER:  35,283
;      REFERENCE/DOCKET NUMBER:  07334/068001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  540 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-019-942-1


  Query Match            99.6%;  Score 2817;  DB 3;  Length 540;
  Best Local Similarity   99.6%;  Pred. No. 1.5e-260;
  Matches  538;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 MNGEAICSALPTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      1 MNGEAICSALPTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSER 60

Qy     61 KDVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPL 120
          |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db     61 KDVLREAEILHKARFSYILPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPL 120

Qy    121 RFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRS 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 RFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRS 180

Qy    181 SKSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMY 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 SKSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMY 240

Qy    241 SVSQGHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEI 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 SVSQGHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEI 300

Qy    301 TFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLNIPVNHGPQEESCGSSQLHENSGSPET 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLNIPVNHGPQEESCGSSQLHENSGSPET 360

Qy    361 SRSLPAPQDNDFLSRKAQDCYFMKLHHCPGNHSWDSTISGSQRAAFCDHKTTPCSSAIIN 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db    361 SRSLPAPQDNDFLSRKAQDCYFMKLHHCPGNHSWDSTISGSQRAAFCDHKTIPCSSAIIN 420

Qy    421 PLSTAGNSERLQPGIAQQWIQSKREDIVNQMTEACLNQSLDALLSRDLIMKEDYELVSTK 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 PLSTAGNSERLQPGIAQQWIQSKREDIVNQMTEACLNQSLDALLSRDLIMKEDYELVSTK 480

Qy    481 PTRTSKVRQLLDTTDIQGEEFAKVIVQKLKDNKQMGLQPYPEILVVSRSPSLNLLQNKSM 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 PTRTSKVRQLLDTTDIQGEEFAKVIVQKLKDNKQMGLQPYPEILVVSRSPSLNLLQNKSM 540

RESULT   2
US-09-188-930-334
; Sequence 334, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
;  APPLICANT: Watson, James D.
;  APPLICANT: Strachan, Lorna
;  APPLICANT: Sleeman, Matthew
;  APPLICANT: Onrust, Rene
;  APPLICANT: Murison, James Greg
;  TITLE OF INVENTION: Compositions Isolated From Skin Cells
;  TITLE OF INVENTION: and Methods For Their Use
;  FILE REFERENCE: 11000.1011c1
;  CURRENT APPLICATION NUMBER: US/09/188,930A
;  CURRENT FILING DATE: 1998-11-09
;  NUMBER OF SEQ ID NOS: 348
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 334
;   LENGTH: 787
;   TYPE: PRT
;   ORGANISM: Mouse
US-09-188-930-334

  Query Match            20.4%;  Score 576;  DB 4;  Length 787;
  Best Local Similarity  33.6%;  Pred. No. 2.4e-46;
  Matches  169;  Conservative   64;  Mismatches  154;  Indels  116;  Gaps   20;

Qy     10 LPTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVK---HLHIHTPLLDSERKDVLRE 66
          | |    : |    :  |  | |   ||  |:  :|:|    ||:     | || ::| |
Db     15 LRTFDAGEFAGWEKVGSGGFGQVYKVRHVHWKTWLAIKCSPSLHVD----DRERMELLEE 70

Qy     67 AEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILH 126
          |: :  |:| || |:  ||| ||  :|:| |||  ||| :||  :     : | |||||:|
Db     71 AKKMEMAKFRYILPVYGICQEP--VGLVMEYMETGSLEKLLASE----PLPWDLRFRIVH 124

Qy    127 EIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPE 186
          | |:|:|:|| |:||||| |||  |||||  :||||:||||:|   || |   |     :
```

```
Db    125 ETAVGMNFLHCMSPPLLHLDLKPANILLDAHYHVKISDFGLAKCNGMSHSHDLSM----D 180

Qy    187 G--GTIIYMPPENYEPGQKSRA-SIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVS 243
          |  ||| |:|||     :|||    |||:||:|:: | ||::|:|| |  | | ||  |
Db    181 GLFGTIAYLPPERIR--EKSRLFDTKHDVYSFAIVIWGVLTQKKPFADEKNILHIMMKVV 238

Qy    244 QGHRPVINEESLPYDIPHR---ARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEI 300
          :||||    |  |   |     | :| |::  |  :|  ||              ||:||
Db    239 KGHRP----ELPPICRPRPRACASLIGLMQRCWHADPQVRP--------------TFQEI 280

Qy    301 TFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLNIPVNHGPQEESCGSSQLHENSGSPET 360
          |                |    ||:|   |:     : | | |:|   |:      |
Db    281 T----------------SETEDLCEKPDEEVK---DLAHEPGEKSSLESKSEARPESSRL 321

Qy    361 SRSLPAPQDNDF----------------------LSRKAQDCYFMKL------HHCPGNH 392
           |:   | |||                       ||| : :|   ||          |
Db    322 KRASAPPFDNDCSLSELLSQLDSGISQTLEGPEELSRSSSEC---KLPSSSSGKRLSGVS 378

Qy    393 SWDSTIS--GSQRAAF-----------CDHKTTPCSSAIINPLSTAGNSERLQPGIAQQW 439
          | ||  |  ||   :|            | :      |||     :|:: ||      :
Db    379 SVDSAFSSRGSLSLSFEREASTGDLGPTDIQKKKLVDAII-----SGDTSRL-----MKI 428

Qy    440 IQSKREDIVNQMTEACLNQSLDA 462
          :| :  |:|   :  : |: :::|
Db    429 LQPQDVDLVLDSSASLLHLAVEA 451

RESULT   3
US-09-188-930-185
; Sequence 185, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
;  APPLICANT: Watson, James D.
;  APPLICANT: Strachan, Lorna
;  APPLICANT: Sleeman, Matthew
;  APPLICANT: Onrust, Rene
;  APPLICANT: Murison, James Greg
;  TITLE OF INVENTION: Compositions Isolated From Skin Cells
;  TITLE OF INVENTION: and Methods For Their Use
;  FILE REFERENCE: 11000.1011c1
;  CURRENT APPLICATION NUMBER: US/09/188,930A
;  CURRENT FILING DATE: 1998-11-09
;  NUMBER OF SEQ ID NOS: 348
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 185
;   LENGTH: 536
;   TYPE: PRT
;   ORGANISM: mouse
US-09-188-930-185

  Query Match            18.9%;  Score 536;  DB 4;  Length 536;
  Best Local Similarity  32.5%;  Pred. No. 8.9e-43;
  Matches  164;  Conservative   64;  Mismatches  159;  Indels  118;  Gaps   21;

Qy     10 LPTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVK---HLHIHTPLLDSERKDVLRE 66
          | |    : |    :  |  | |   ||  |:  :|:|    ||:     | || ::| |
Db     15 LRTFDAGEFAGWEKVGSGGFGQVYKVRHVHWKTWLAIKCSPSLHVD----DRERMELLEE 70

Qy     67 AEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILH 126
          |: :  |:| || |:  ||| ||  :|:| |||  ||| :||  :     : | |||||:|
Db     71 AKKMEMAKFRYILPVYGICQEP--VGLVMEYMETGSLEKLLASE----PLPWDLRFRIVH 124

Qy    127 EIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHV-KIADFGLSKWRMMSLSQSRSSKSAP 185
          | |:|:|:|| |:||||| |||  |||||  : : :  ||||:|   || |   |
Db    125 ETAVGMNFLHCMSPPLLHLDLKPANILLDAHYQMSRFLDFGLAKCNGMSHSHDLSM---- 180

Qy    186 EG--GTIIYMPPENYEPGQKSRA-SIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSV 242
          :|  ||| |:|||     :|||    |||:||:|:: | ||::  || |  | | ||  |
Db    181 DGLFGTIGYLPPERIR--EKSRLFDTKHDVYSFAIVIWGVLTQNNPFADEKNILHIMMKV 238

Qy    243 SQGHRPVINEESLPYDIPHR---ARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEE 299
           :||||    |  |   |     | :| |::  |  :|  ||              ||:|
```

```
Db    239 VKGHRP----ELPPICRPRPRACASLIGLMQRCWHADPQVRP--------------TFQE 280

Qy    300 ITFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLNIPVNHGPQEESCGSSQLHENSGSPE 359
          ||                |    ||:|   |:     : | | |:|   |:      |
Db    281 IT----------------SETEDLCEKPDEEVK---DLAHEPGEKSSLESKSEARPESSR 321

Qy    360 TSRSLPAPQDNDF-----------------------LSRKAQDCYFMKL------HHCPG 390
            |:   | |||                        ||| : :|   ||          |
Db    322 LKRASAPPFDNDCSLSELLSQLDSGIFPRLLKGPEELSRSSSEC---KLPSSSSGKRLSG 378

Qy    391 NHSWDSTIS--GSQRAAF-----------CDHKTTPCSSAIINPLSTAGNSERLQPGIAQ 437
            | ||  |  ||   :|            | :      |||     :|:: ||
Db    379 VSSVDSAFSSRGSLSLSFEREASTGDLGPTDIQKKKLVDAII-----SGDTSRL-----M 428

Qy    438 QWIQSKREDIVNQMTEACLNQSLDA 462
          : :| :  |:|   : : |: :::|
Db    429 KILQPQDVDLVLDSSASLLHLAVEA 453

RESULT   4
US-08-444-005-15
; Sequence 15, Application US/08444005
; Patent No. 5674734
;  GENERAL INFORMATION:
;    APPLICANT:  Leder, Philip
;    APPLICANT:  Seed, Brian
;    APPLICANT:  Stanger, Ben Z.
;    APPLICANT:  Lee, Tae-Ho
;    APPLICANT:  Kim, Emily
;    TITLE OF INVENTION:  CELL DEATH PROTEIN
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street, Suite 3100
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/444,005
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,164
;      REFERENCE/DOCKET NUMBER:  00383/026001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  656 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-444-005-15


  Query Match              13.9%;  Score 394;  DB 1;  Length 656;
  Best Local Similarity    31.0%;  Pred. No. 4.4e-29;
  Matches  131;  Conservative   66;  Mismatches  144;  Indels   82;  Gaps   19;

Qy     24 LSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSERKDV-LREAEILHKARFSYIFPIL 82
          |  |  | ||   |      :  |   ::|    :|  :| | | :::|: | | :  :|
Db     23 LDSGGFGKVSLCYHRSHGFVILKK---VYTGPNRAEYNEVLLEEGKMMHRLRHSRVVKLL 79
```

```
Qy     83 GICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPL 142
          ||  |     :| |||  |:|  :|  ||:  ||   |: ||: |   |: |||:    :
Db     80 GIIIEEGNYSLVMEYMEKGNLMHVL--KTQI-DVPLSLKGRIIVEAIEGMCYLHD--KGV 134

Qy    143 LHHDLKTQNILLDNEFHVKIADFGLSKWRMMS-LSQSRSSKSAP--------EGGTIIYM 193
          :| ||| :|||:| :||:|||| |:: ::  | |:: : :|            |||: ||
Db    135 IHKDLKPENILVDRDFHIKIADLGVASFKTWSKLTKEKDNKQKEVSSTTKKNNGGTLYYM 194

Qy    194 PPE-----NYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVSQGHRP 248
           ||     | :| :||      |:||: :: | : ::|:|:|:|    | :  :  |:||
Db    195 APEHLNDINAKPTEKS------DVYSFGIVLWAIFAKKEPYENVICTEQFVICIKSGNRP 248

Qy    249 VINE--ESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEP-VLRTFEEITFLEA 305
           : |  |  | :|      |||:|  |   |::||:||    |  |  |  |||    |
Db    249 NVEEILEYCPREI------ISLMERCWQAIPEDRPTFLGIEEEFRPFYLSHFEEYV-EED 301

Qy    306 VIQLKK------TKLQSVSSAIHLC---------DKKKMELSLNIPVNHGPQEESCGSSQ 350
          |  |||        || : |  | |          ::   |  :  :  || |||  ||
Db    302 VASLKKEYPDQSPVLQRMFSLQHDCVPLPPSRSNSEQPGSLHSSQGLQMGPVEESWFSS- 360

Qy    351 LHENSGSPETSRSLPAPQDNDFLSRKAQDCYFMKLHHCPGNHSWDSTISGSQRAAFCDHK 410
                |||       ||| :  | :|      ||      |::           | : :
Db    361 ------SPE------YPQDENDRSVQA------KLQEEASYHAF---------GIFAEKQ 393

Qy    411 TTP 413
          | |
Db    394 TKP 396

RESULT   5
US-09-329-418-9
; Sequence 9, Application US/09329418
; Patent No. 6096539
; GENERAL INFORMATION:
;  APPLICANT: ZENECA Limited
;  TITLE OF INVENTION: PROTEIN ACTIVATOR OF APOPTOSIS
;  FILE REFERENCE: PHM.70536
;  CURRENT APPLICATION NUMBER: US/09/329,418
;  CURRENT FILING DATE: 1999-06-11
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 9
;   LENGTH: 518
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Dominant Negative Mutant Embodiment
US-09-329-418-9


  Query Match             13.8%;  Score 389;  DB 3;  Length 518;
  Best Local Similarity   25.7%;  Pred. No. 9.1e-29;
  Matches  132;  Conservative   84;  Mismatches  183;  Indels  114;  Gaps   20;

Qy     11 PTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSERKDVLREAEIL 70
          | :   :| :   : :|  |||  |:|  |   |||| ::          | : || : :
Db     14 PLVSIEELENQELVGKGGFGTVFRAQHRKWGYDVAVKIVN---------SKAISREVKAM 64

Qy     71 HKARFSYIFPILGICNEPEF-----LGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRIL 125
                ::  : |:  :  :       :||::| ||||: ||  :   |   |||  |:|
Db     65 ASLDNEFVLRLEGVIEKVNWDQDPKPALVTKFMENGSLSGLLQSQCPRP---WPLLCRLL 121

Qy    126 HEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAP 185
           |: ||: |||:  | ||| |||  |:||| | |||:|||||| ::    ||| :    |
Db    122 KEVVLGMFYLHDQNPVLLHRDLKPSNVLLDPELHVKLADFGLSTFQ--GGSQSGTGSGEP 179

Qy    186 EGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSV-SQ 244
           |||: |: || :      :||   |:||: :: | ||: ::  |  | |  :  :| ::
Db    180 -GGTLGYLAPELF-VNVNRKASTASDVYSFGILMWAVLAGRE-VELPTEPSLVYEAVCNR 236

Qy    245 GHRPVINEESLPY---DIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEE-- 299
           :|| : |  ||    : |    :  |::  |:  | :|||| :|| : : | :  |
Db    237 QNRPSLAE--LPQAGPETPGLEGLKELMQLCWSSEPKDRPSFQECLPKTDEVFQMVENNM 294
```

```
Qy     300 ----ITFLEAVIQLKK-------------TKLQSVSSAIH--------LCDKKKMELSL 333
                |  : : ||:               |::      |         :  :   :|:|
Db     295 NAAVSTVKDFLSQLRSSNRRFSIPESGQGGTEMDGFRRTIENQHSRNDVMVSEWLNKLNL 354

Qy     334 NIPVNHGP-----------QEE--------SCGSSQLHENSGSPETS------------ 361
             | :  |            |||           |  : :   :||||
Db     355 EEPPSSVPKKCPSLTKRSRAQEEQVPQAWTAGTSSDSMAQPPQTPETSTFRNQMPSPTST 414

Qy     362 -RSLPAPQDNDFLSRKAQDCY------------FMKLHHCPGNHSWDSTISGSQRAAFC 407
               | |: |    |:  :                : :::| |    |:     |
Db     415 GTPSPGPRGNQGAERQGMNASCRTPEPNPVTGRPLVNIYNCSGVQVGDNNYLTMQ----- 469

Qy     408 DHKTTPCSSAIINPLSTAGNSERLQ--PGIAQQ 438
             :||   :| : |   :|    ||  | :  |
Db     470 --QTTALPTAGLAP---SGKGRGLQHPPPVGSQ 497

RESULT   6
US-09-329-418-3
; Sequence 3, Application US/09329418
; Patent No. 6096539
; GENERAL INFORMATION:
;  APPLICANT: ZENECA Limited
;  TITLE OF INVENTION: PROTEIN ACTIVATOR OF APOPTOSIS
;  FILE REFERENCE: PHM.70536
;  CURRENT APPLICATION NUMBER: US/09/329,418
;  CURRENT FILING DATE: 1999-06-11
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 518
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-09-329-418-3


  Query Match             13.7%;  Score 388;  DB 3;  Length 518;
  Best Local Similarity   32.0%;  Pred. No. 1.1e-28;
  Matches  102;  Conservative   62;  Mismatches  125;  Indels   30;  Gaps   11;

Qy      11 PTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSERKDVLREAEIL 70
           | :   :| :   : :|  |||  |:|  |   |||| ::          | : || : :
Db      14 PLVSIEELENQELVGKGGFGTVFRAQHRKWGYDVAVKIVN---------SKAISREVKAM 64

Qy      71 HKARFSYIFPILGICNEPEF-----LGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRIL 125
                 ::  : |:  :  :       :||::| ||||: ||  :   |   |||  |:|
Db      65 ASLDNEFVLRLEGVIEKVNWDQDPKPALVTKFMENGSLSGLLQSQCPRP---WPLLCRLL 121

Qy     126 HEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAP 185
            |: ||: |||:  | ||| |||  |:||| | |||:|||||| ::    ||| :    |
Db     122 KEVVLGMFYLHDQNPVLLHRDLKPSNVLLDPELHVKLADFGLSTFQ--GGSQSGTGSGEP 179

Qy     186 EGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSV-SQ 244
            |||: |: || :      :||   |:||: :: | ||: ::  |  | |  :  :| ::
Db     180 -GGTLGYLAPELF-VNVNRKASTASDVYSFGILMWAVLAGRE-VELPTEPSLVYEAVCNR 236

Qy     245 GHRPVINEESLPY---DIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEIT 301
            :|| : |  ||    : |    :  |::  |:  | :|||| :|| : : | :  |
Db     237 QNRPSLAE--LPQAGPETPGLEGLKELMQLCWSSEPKDRPSFQECLPKTDEVFQMVE--N 292

Qy     302 FLEAVIQLKKTKLQSVSSA 320
            : | :   |  |  : |:
Db     293 NMNAAVSTVKDFLSQLRSS 311

RESULT   7
US-09-329-418-8
; Sequence 8, Application US/09329418
; Patent No. 6096539
; GENERAL INFORMATION:
;  APPLICANT: ZENECA Limited
;  TITLE OF INVENTION: PROTEIN ACTIVATOR OF APOPTOSIS
;  FILE REFERENCE: PHM.70536
;  CURRENT APPLICATION NUMBER: US/09/329,418
;  CURRENT FILING DATE: 1999-06-11
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
;   LENGTH: 420
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Delta Death Domain
US-09-329-418-8

  Query Match             13.7%;  Score 386.5;  DB 3;  Length 420;
  Best Local Similarity   32.1%;  Pred. No. 1.1e-28;
  Matches  103;  Conservative   62;  Mismatches  123;  Indels   33;  Gaps   12;

Qy      11 PTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSERKDVLREAEIL 70
           | :   :| :   : :|  |||  |:|  |   |||| ::          | : || : :
Db      14 PLVSIEELENQELVGKGGFGTVFRAQHRKWGYDVAVKIVN---------SKAISREVKAM 64

Qy      71 HKARFSYIFPILGI-------CNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFR 123
                 ::  : |:         :|:   :||::| ||||: ||  :   |   |||  |
Db      65 ASLDNEFVLRLEGVIEKVGGSSQDPK-PALVTKFMENGSLSGLLQSQCPRP---WPLLCR 120

Qy     124 ILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKS 183
           :| |: ||: |||:  | ||| |||  |:||| | |||:|||||| ::    ||| :
Db     121 LLKEVVLGMFYLHDQNPVLLHRDLKPSNVLLDPELHVKLADFGLSTFQ--GGSQSGTGSG 178

Qy     184 APEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSV- 242
            | |||: |: || :      :||   |:||: :: | ||: ::  |  | |  :  :|
Db     179 EP-GGTLGYLAPELF-VNVNRKASTASDVYSFGILMWAVLAGRE-VELPTEPSLVYEAVC 235

Qy     243 SQGHRPVINEESLPY---DIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEE 299
           :: :|| : |  ||    : |    :  |::  |:  | :|||| :|| : : | :  |
Db     236 NRQNRPSLAE--LPQAGPETPGLEGLKELMQLCWSSEPKDRPSFQECLPKTDEVFQMVE- 292

Qy     300 ITFLEAVIQLKKTKLQSVSSA 320
              : | :   |  |  : |:
Db     293 -NNMNAAVSTVKDFLSQLRSS 312

RESULT   8
US-09-329-418-5
; Sequence 5, Application US/09329418
; Patent No. 6096539
; GENERAL INFORMATION:
;  APPLICANT: ZENECA Limited
;  TITLE OF INVENTION: PROTEIN ACTIVATOR OF APOPTOSIS
;  FILE REFERENCE: PHM.70536
;  CURRENT APPLICATION NUMBER: US/09/329,418
;  CURRENT FILING DATE: 1999-06-11
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
;   LENGTH: 518
;   TYPE: PRT
;   ORGANISM: DArtificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Dominant Negative Mutant Embodiment
US-09-329-418-5

  Query Match             13.6%;  Score 385;  DB 3;  Length 518;
  Best Local Similarity   31.7%;  Pred. No. 2.2e-28;
  Matches  101;  Conservative   63;  Mismatches  125;  Indels   30;  Gaps   11;

Qy      11 PTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSERKDVLREAEIL 70
           | :   :| :   : :|  |||  |:|  |   |||: ::          | : || : :
Db      14 PLVSIEELENQELVGKGGFGTVFRAQHRKWGYDVAVRIVN---------SKAISREVKAM 64

Qy      71 HKARFSYIFPILGICNEPEF-----LGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRIL 125
```

```
                 ::  : |:  :  :       :||::| ||||: ||  :   |   |||  |:|
Db       65 ASLDNEFVLRLEGVIEKVNWDQDPKPALVTKFMENGSLSGLLQSQCPRP---WPLLCRLL 121

Qy      126 HEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAP 185
             |: ||: |||:  | ||| |||  |:||| | |||:|||||| ::    ||| :    |
Db      122 KEVVLGMFYLHDQNPVLLHRDLKPSNVLLDPELHVKLADFGLSTFQ--GGSQSGTGSGEP 179

Qy      186 EGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSV-SQ 244
             |||: |: || :      :||   |:||: :: | ||: ::  |  | |  :  :| ::
Db      180 -GGTLGYLAPELF-VNVNRKASTASDVYSFGILMWAVLAGRE-VELPTEPSLVYEAVCNR 236

Qy      245 GHRPVINEESLPY---DIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEIT 301
             :|| : |  ||    : |    :  |::  |:  | :|||| :|| : : | :  |
Db      237 QNRPSLAE--LPQAGPETPGLEGLKELMQLCWSSEPKDRPSFQECLPKTDEVFQMVE--N 292

Qy      302 FLEAVIQLKKTKLQSVSSA 320
             : | :   |  |  : |:
Db      293 NMNAAVSTVKDFLSQLRSS 311


RESULT   9
US-09-329-418-4
; Sequence 4, Application US/09329418
; Patent No. 6096539
; GENERAL INFORMATION:
;  APPLICANT: ZENECA Limited
;  TITLE OF INVENTION: PROTEIN ACTIVATOR OF APOPTOSIS
;  FILE REFERENCE: PHM.70536
;  CURRENT APPLICATION NUMBER: US/09/329,418
;  CURRENT FILING DATE: 1999-06-11
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 518
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Dominant Negative Mutant Embodiment
US-09-329-418-4

  Query Match             13.4%;  Score 380;  DB 3;  Length 518;
  Best Local Similarity   31.7%;  Pred. No. 6.6e-28;
  Matches  101;  Conservative   62;  Mismatches  126;  Indels   30;  Gaps   11;

Qy       11 PTIPYHKLADLRYLSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSERKDVLREAEIL 70
            | :   :| :   : :|  |||  |:|  |   |||| ::          | : || : :
Db       14 PLVSIEELENQELVGKGGFGTVFRAQHRKWGYDVAVKIVN---------SKAISREVKAM 64

Qy       71 HKARFSYIFPILGICNEPEF-----LGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRIL 125
                  ::  : |:  :  :       :||::| ||||: ||  :   |   |||  |:|
Db       65 ASLDNEFVLRLEGVIEKVNWDQDPKPALVTKFMENGSLSGLLQSQCPRP---WPLLCRLL 121

Qy      126 HEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAP 185
             |: ||: |||:  | |||  ||  |:||| | |||:|||||| ::    ||| :    |
Db      122 KEVVLGMFYLHDQNPVLLHRALKPSNVLLDPELHVKLADFGLSTFQ--GGSQSGTGSGEP 179

Qy      186 EGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSV-SQ 244
             |||: |: || :      :||   |:||: :: | ||: ::  |  | |  :  :| ::
Db      180 -GGTLGYLAPELF-VNVNRKASTASDVYSFGILMWAVLAGRE-VELPTEPSLVYEAVCNR 236

Qy      245 GHRPVINEESLPY---DIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEIT 301
             :|| : |  ||    : |    :  |::  |:  | :|||| :|| : : | :  |
Db      237 QNRPSLAE--LPQAGPETPGLEGLKELMQLCWSSEPKDRPSFQECLPKTDEVFQMVE--N 292

Qy      302 FLEAVIQLKKTKLQSVSSA 320
             : | :   |  |  : |:
Db      293 NMNAAVSTVKDFLSQLRSS 311


RESULT  10
US-07-928-464-2
; Sequence 2, Application US/07928464
; Patent No. 5367065
;  GENERAL INFORMATION:
;    APPLICANT:  Ecker, Joseph R.
;    APPLICANT:  Kieber, Joseph J.
;    TITLE OF INVENTION:  Constitutive Triple Response Gene and
;    TITLE OF INVENTION:  Mutations
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock, Washburn, Kurtz, Mackiewicz and
;      ADDRESSEE:  No. 5367065ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  U.S.A.
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/928,464
;      FILING DATE:  19920810
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Miller, Suzanne E.
;      REGISTRATION NUMBER:  32,279
;      REFERENCE/DOCKET NUMBER:  UPN-1086
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  821 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-928-464-2

  Query Match             13.3%;  Score 376;  DB 1;  Length 821;
  Best Local Similarity   30.7%;  Pred. No. 3.2e-27;
  Matches   98;  Conservative   60;  Mismatches  111;  Indels   50;  Gaps   12;

Qy        1 MNGEAICSALPT--------------IPYHKLADLRYLSRGASGTVSSARHADWR-VQVA 45
            ||   |   :|               ||:    ||    :  :|:  :   |:|    ||
Db      520 MNAPPISQPVPNRANRELGLDGDDMDIPW---CDLNIKEKIGAGSFGTVHRAEWHGSDVA 576

Qy       46 VKHL---HIHTPLLDSER-KDVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNG 101
            || |     |     :||  : |||  |: : |   |   :|   :|  | |||||:  |
Db      577 VKILMEQDFH-----AERVNEFLREVAIMKRLRHPNIVLFMGAVTQPPNLSIVTEYLSRG 631

Qy      102 SLNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVK 161
            ||  |||:      :    |  : :::| |:|||||  ||::| |||: |:|:| :: ||
Db      632 SLYRLLHKSGAREQLDERRRLSMAYDVAKGMNYLHNRNPPIVHRDLKSPNLLVDKKYTVK 691

Qy      162 IADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPE--NYEPGQKSRASIKHDIYSYAVIT 219
            : |||||: :  :   |:|:   ||     :| ||    ||  :     | |:||: ||
Db      692 VCDFGLSRLKASTFLSSKSAAGTPE-----WMAPEVLRDEPSNE-----KSDVYSFGVIL 741

Qy      220 WEVLSRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIPH--RARMISLIESGWAQNP 277
            ||: : :||: :: || |:: :|    : :        :||     :: ::||  |   |
Db      742 WELATLQQPWGNL-NPAQVVAAVGFKCKRL--------EIPRNLNPQVAAIIEGCWTNEP 792

Qy      278 DERPSFLKCLIELEPVLRT 296
             :||||   :  | |::::
Db      793 WKRPSFATIMDLLRPLIKS 811


RESULT  11
US-08-003-311B-2
; Sequence 2, Application US/08003311B
; Patent No. 5444166
;  GENERAL INFORMATION:
```

```
;   APPLICANT:  Ecker, Joseph R.
;   APPLICANT:  Kieber, Joseph J.
;   TITLE OF INVENTION:  Constitutive Triple Response Gene
;   TITLE OF INVENTION:  and Mutations
;   NUMBER OF SEQUENCES:  13
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz and
;     ADDRESSEE:  No. 5444166ris
;     STREET:  One Liberty Place - 46th Floor
;     CITY:  Philadelphia
;     STATE:  PA
;     COUNTRY:  U.S.A.
;     ZIP:  19103
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  WordPerfect 5.1
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/003,311B
;     FILING DATE:  January 12, 1993
;     CLASSIFICATION:  800
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  07/928,464
;     FILING DATE:  August 10, 1992
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Lori Y. Beardell
;     REGISTRATION NUMBER:  34,293
;     REFERENCE/DOCKET NUMBER:  UPN-1108
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  215-568-3100
;     TELEFAX:  215-568-3439
; INFORMATION FOR SEQ ID NO:  2:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  821 amino acids
;     TYPE:  amino acid
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  protein
US-08-003-311B-2

  Query Match             13.3%;  Score 376;  DB 1;  Length 821;
  Best Local Similarity   30.7%;  Pred. No. 3.2e-27;
  Matches   98;  Conservative   60;  Mismatches  111;  Indels   50;  Gaps   12;

Qy      1 MNGEAICSALPT--------------IPYHKLADLRYLSRGASGTVSSARHADWR-VQVA 45
          ||   |   :|               ||:    ||    :  :|:  :   |:|    ||
Db    520 MNAPPISQPVPNRANRELGLDGDDMDIPW---CDLNIKEKIGAGSFGTVHRAEWHGSDVA 576

Qy     46 VKHL---HIHTPLLDSER-KDVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNG 101
          || |     |     :||  : |||  |: : |   |   :|   :|  | |||||:  |
Db    577 VKILMEQDFH-----AERVNEFLREVAIMKRLRHPNIVLFMGAVTQPPNLSIVTEYLSRG 631

Qy    102 SLNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVK 161
          ||  |||:      :    |  : :::| |:|||||  ||::| |||: |:|:| :: ||
Db    632 SLYRLLHKSGAREQLDERRRLSMAYDVAKGMNYLHNRNPPIVHRDLKSPNLLVDKKYTVK 691

Qy    162 IADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPE--NYEPGQKSRASIKHDIYSYAVIT 219
          : |||||: :  :   |:|:   ||     :| ||    ||  :     | |:||: ||
Db    692 VCDFGLSRLKASTFLSSKSAAGTPE-----WMAPEVLRDEPSNE-----KSDVYSFGVIL 741

Qy    220 WEVLSRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIPH--RARMISLIESGWAQNP 277
          ||: : :||: :: || |:: :|    : :        :||     :: ::||  |   |
Db    742 WELATLQQPWGNL-NPAQVVAAVGFKCKRL--------EIPRNLNPQVAAIIEGCWTNEP 792

Qy    278 DERPSFLKCLIELEPVLRT 296
           :||||   :  | |::::
Db    793 WKRPSFATIMDLLRPLIKS 811

RESULT 12
US-08-261-432-2
; Sequence 2, Application US/08261432
```

```
; Patent No. 5602322
; GENERAL INFORMATION:
;   APPLICANT:  Ecker, Joseph R.
;   APPLICANT:  Kieber, Joseph J.
;   TITLE OF INVENTION:  Constitutive Triple Response Gene
;   TITLE OF INVENTION:  and Mutations
;   NUMBER OF SEQUENCES:  13
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz and
;     ADDRESSEE:  No. 5602322ris
;     STREET:  One Liberty Place - 46th Floor
;     CITY:  Philadelphia
;     STATE:  PA
;     COUNTRY:  U.S.A.
;     ZIP:  19103
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  WordPerfect 5.1
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/261,432
;     FILING DATE:  June 17, 1994
;     CLASSIFICATION:  800
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  08/003,311
;     FILING DATE:  January 12, 1993
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Lori Y. Beardell
;     REGISTRATION NUMBER:  34,293
;     REFERENCE/DOCKET NUMBER:  UPN-1864
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  215-568-3100
;     TELEFAX:  215-568-3439
; INFORMATION FOR SEQ ID NO:  2:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  821 amino acids
;     TYPE:  amino acid
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  protein
US-08-261-432-2

  Query Match             13.3%;  Score 376;  DB 1;  Length 821;
  Best Local Similarity   30.7%;  Pred. No. 3.2e-27;
  Matches   98;  Conservative   60;  Mismatches  111;  Indels   50;  Gaps   12;

Qy      1 MNGEAICSALPT--------------IPYHKLADLRYLSRGASGTVSSARHADWR-VQVA 45
          ||   |   :|               ||:    ||    :  :|:  :   |:|    ||
Db    520 MNAPPISQPVPNRANRELGLDGDDMDIPW---CDLNIKEKIGAGSFGTVHRAEWHGSDVA 576

Qy     46 VKHL---HIHTPLLDSER-KDVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNG 101
          || |     |     :||  : |||  |: : |   |   :|   :|  | |||||:  |
Db    577 VKILMEQDFH-----AERVNEFLREVAIMKRLRHPNIVLFMGAVTQPPNLSIVTEYLSRG 631

Qy    102 SLNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVK 161
          ||  |||:      :    |  : :::| |:|||||  ||::| |||: |:|:| :: ||
Db    632 SLYRLLHKSGAREQLDERRRLSMAYDVAKGMNYLHNRNPPIVHRDLKSPNLLVDKKYTVK 691

Qy    162 IADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPE--NYEPGQKSRASIKHDIYSYAVIT 219
          : |||||: :  :   |:|:   ||     :| ||    ||  :     | |:||: ||
Db    692 VCDFGLSRLKASTFLSSKSAAGTPE-----WMAPEVLRDEPSNE-----KSDVYSFGVIL 741

Qy    220 WEVLSRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIPH--RARMISLIESGWAQNP 277
          ||: : :||: :: || |:: :|    : :        :||     :: ::||  |   |
Db    742 WELATLQQPWGNL-NPAQVVAAVGFKCKRL--------EIPRNLNPQVAAIIEGCWTNEP 792

Qy    278 DERPSFLKCLIELEPVLRT 296
           :||||   :  | |::::
Db    793 WKRPSFATIMDLLRPLIKS 811

RESULT 13
```

```
PCT-US93-07347-2
; Sequence 2, Application PC/TUS9307347
;  GENERAL INFORMATION:
;    APPLICANT:  Ecker, Joseph R.
;    APPLICANT:  Kieber, Joseph J.
;    TITLE OF INVENTION:  Constitutive Triple Response Gene and
;    TITLE OF INVENTION:  Mutations
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock, Washburn, Kurtz, Mackiewicz and
;      ADDRESSEE:  Norris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  U.S.A.
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/07347
;      FILING DATE:  19930805
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Miller, Suzanne E.
;      REGISTRATION NUMBER:  32,279
;      REFERENCE/DOCKET NUMBER:  UPN-1086
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  821 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US93-07347-2


  Query Match             13.3%;  Score 376;  DB 5;  Length 821;
  Best Local Similarity   30.7%;  Pred. No. 3.2e-27;
  Matches   98;  Conservative   60;  Mismatches  111;  Indels   50;  Gaps   12;

Qy          1 MNGEAICSALPT--------------IPYHKLADLRYLSRGASGTVSSARHADWR-VQVA 45
              ||   |   :|               ||:    ||    :  :|:  :   |:|    ||
Db        520 MNAPPISQPVPNRANRELGLDGDDMDIPW---CDLNIKEKIGAGSFGTVHRAEWHGSDVA 576

Qy         46 VKHL---HIHTPLLDSER-KDVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNG 101
              || |     |     :||  : ||| |:  : |   |   :|   :|  | |||||:  |
Db        577 VKILMEQDFH-----AERVNEFLREVAIMKRLRHPNIVLFMGAVTQPPNLSIVTEYLSRG 631

Qy        102 SLNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVK 161
              ||  |||:      :    |  : :::| |:|||||  ||::| |||: |:|:| :: ||
Db        632 SLYRLLHKSGAREQLDERRRLSMAYDVAKGMNYLHNRNPPIVHRDLKSPNLLVDKKYTVK 691

Qy        162 IADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPE--NYEPGQKSRASIKHDIYSYAVIT 219
              : |||||: :  :   |:|:   ||     :| ||    ||  :     | |:||: ||
Db        692 VCDFGLSRLKASTFLSSKSAAGTPE-----WMAPEVLRDEPSNE-----KSDVYSFGVIL 741

Qy        220 WEVLSRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIPH--RARMISLIESGWAQNP 277
              ||: : :||: :: || |:: :|    : :        :||     :: ::||  |   |
Db        742 WELATLQQPWGNL-NPAQVVAAVGFKCKRL--------EIPRNLNPQVAAIIEGCWTNEP 792

Qy        278 DERPSFLKCLIELEPVLRT 296
               :||||   :  | |::::
Db        793 WKRPSFATIMDLLRPLIKS 811

RESULT  14
US-09-329-418-6
; Sequence 6, Application US/09329418
; Patent No. 6096539
```

```
; GENERAL INFORMATION:
;  APPLICANT: ZENECA Limited
;  TITLE OF INVENTION: PROTEIN ACTIVATOR OF APOPTOSIS
;  FILE REFERENCE: PHM.70536
;  CURRENT APPLICATION NUMBER: US/09/329,418
;  CURRENT FILING DATE: 1999-06-11
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
;   LENGTH: 261
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Kinase Domain
US-09-329-418-6

  Query Match             13.1%;  Score 372;  DB 3;  Length 261;
  Best Local Similarity   34.1%;  Pred. No. 1.3e-27;
  Matches   95;  Conservative   53;  Mismatches  103;  Indels   28;  Gaps   10;

Qy         18 LADLRYLSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSERKDVLREAEILHKARFSY 77
              | :   : :|  |||  |:|  |   |||| ::          | : || : :      :
Db          1 LENQELVGKGGFGTVFRAQHRKWGYDVAVKIVN---------SKAISREVKAMASLDNEF 51

Qy         78 IFPILGICNEPEF-----LGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILHEIALGV 132
              :  : |:  :  :       :||::| ||||: ||  :   |   |||  |:| |: ||:
Db         52 VLRLEGVIEKVNWDQDPKPALVTKFMENGSLSGLLQSQCPRP---WPLLCRLLKEVVLGM 108

Qy        133 NYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPEGGTIIY 192
               |||:  | ||| |||  |:||| | |||:|||||| ::    ||| :    | |||: |
Db        109 FYLHDQNPVLLHRDLKPSNVLLDPELHVKLADFGLSTFQ--GGSQSGTGSGEP-GGTLGY 165

Qy        193 MPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSV-SQGHRPVIN 251
              : || :      :||   |:||: :: | ||: ::  |  | |  :  :| :: :|| :
Db        166 LAPELF-VNVNRKASTASDVYSFGILMWAVLAGRE-VELPTEPSLVYEAVCNRQNRPSLA 223

Qy        252 EESLPY---DIPHRARMISLIESGWAQNPDERPSFLKCL 287
              |  ||    : |    :  |::  |:  | :|||| :||
Db        224 E--LPQAGPETPGLEGLKELMQLCWSSEPKDRPSFQECL 260

RESULT  15
US-09-035-706-5
; Sequence 5, Application US/09035706
; Patent No. 6001622
;  GENERAL INFORMATION:
;    APPLICANT:  Dedhar, Shoukat
;    APPLICANT:  Hannigan, Greg
;    TITLE OF INVENTION:  Integrin-Linked Kinase and
;    TITLE OF INVENTION:  its Uses
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Bozicevic & Reed, LLP
;      STREET:  285 Hamilton Avenue, Suite 200
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/035,706
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sherwood, Pamela J
```

```
;      REGISTRATION NUMBER:  36,677
;      REFERENCE/DOCKET NUMBER:  KIN-2CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-327-3400
;      TELEFAX:  650 327-3231
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  263 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-035-706-5

  Query Match             13.0%;  Score 369;  DB 3;  Length 263;
  Best Local Similarity   32.4%;  Pred. No. 2.6e-27;
  Matches   94;  Conservative   56;  Mismatches  104;  Indels   36;  Gaps   11;

Qy      13 IPYHKLADLRYLSRGASGTVSSARHADWR-VQVAVKHL---HIHTPLLDSER-KDVLREA 67
           ||:    ||    :  :|:  :   |:|    |||| |     |     :||  : |||
Db       1 IPW---CDLNIKEKIGAGSFGTVHRAEWHGSDVAVKILMEQDFH-----AERVNEFLREV 52

Qy      68 EILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILHE 127
            |: : |   |   :|   :|  | |||||:  |||  |||:      :    |  : ::
Db      53 AIMKRLRHPNIVLFMGAVTQPPNLSIVTEYLSRGSLYRLLHKSGAREQLDERRRLSMAYD 112

Qy     128 IALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPEG 187
           :| |:|||||  ||::| |||: |:|:| :: ||: |||||: :  :   |:|:   ||
Db     113 VAKGMNYLHNRNPPIVHRDLKSPNLLVDKKYTVKVCDFGLSRLKASTFLSSKSAAGTPE- 171

Qy     188 GTIIYMPPE--NYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVSQG 245
               :| ||    ||  :     | |:||: || ||: : :||: :: || |:: :|
Db     172 ----WMAPEVLRDEPSNE-----KSDVYSFGVILWELATLQQPWGNL-NPAQVVAAVGFK 221

Qy     246 HRPVINEESLPYDIPH--RARMISLIESGWAQNPDERPSFLKCLIELEPV 293
            : :        :||     :: ::||  |   | :||||   :  | |:
Db     222 CKRL--------EIPRNLNPQVAAIIEGCWTNEPWKRPSFATIMDLLRPL 263


Search completed: June 14, 2001, 08:06:04
Job time: 96 sec
```

```
                    GenCore version 4.5
          Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          June 14, 2001, 08:04:28 ; Search time 13.68 Seconds
                                           (without alignments)
                                           1352.190 Million cell updates/sec

Title:           US-09-445-223-1
Perfect score:   2829
Sequence:        1 MNGEAICSALPTIPYHKLAD..........PEILVVSRSPSLNLLQNKSM 540

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters:      93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_39:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                   Description
-------------------------------------------------------------------------------
    1   541.5   19.1    832  1  ANR3_HUMAN           P57078 homo sapien
    2     394   13.9    656  1  RIP_MOUSE            Q60855 mus musculu
    3     376   13.3    821  1  CTR1_ARATH           Q05609 arabidopsis
    4     369   13.0    671  1  RIP_HUMAN            Q13546 homo sapien
    5   313.5   11.1   1584  1  KYK1_DICDI           P18160 dictyosteli
    6     311   11.0    579  1  M3K7_MOUSE           Q62073 mus musculu
    7   303.5   10.7    410  1  KYK2_DICDI           P18161 dictyosteli
    8   300.5   10.6    606  1  M3K7_HUMAN           O43318 homo sapien
    9     295   10.4   1050  1  ANPB_ANGJA           P55202 anguilla ja
   10   292.5   10.3   1047  1  ANPB_RAT             P16067 rattus norv
   11     290   10.3    954  1  M3KA_HUMAN           Q02779 homo sapien
   12   289.5   10.2    999  1  RLK5_ARATH           P47735 arabidopsis
   13   287.5   10.2    813  1  KRAF_CAEEL           Q07292 caenorhabdi
   14   287.5   10.2   1047  1  ANPB_HUMAN           P20594 homo sapien
   15     284   10.0    394  1  M3K9_HUMAN           P80192 homo sapien
   16   280.5    9.9   1047  1  ANPB_BOVIN           P46197 bos taurus
   17   277.5    9.8    328  1  KRAB_MOUSE           P28028 mus musculu
   18   277.5    9.8    367  1  RMIL_AVII1           P10533 avian retro
   19   277.5    9.8    450  1  RMIL_AVEVR           P27966 avian rous-
   20   277.5    9.8    712  1  IRA1_HUMAN           P51617 homo sapien
   21   277.5    9.8    765  1  KRAB_HUMAN           P15056 homo sapien
   22   277.5    9.8    806  1  RMIL_CHICK           Q04982 gallus gall
   23   277.5    9.8    807  1  RMIL_COTJA           P34908 coturnix co
   24   273.5    9.7    527  1  TXK_MOUSE            P42682 mus musculu
   25     273    9.7    618  1  ZA70_MOUSE           P43404 mus musculu
   26   272.5    9.6    630  1  TEC_MOUSE            P24604 mus musculu
   27   272.5    9.6    631  1  TEC_HUMAN            P42680 homo sapien
   28     272    9.6    781  1  KRAF_DROME           P11346 drosophila
   29   271.5    9.6    497  1  SPK1_DUGTI           P42687 dugesia tig
   30     270    9.5    619  1  ZA70_HUMAN           P43403 homo sapien
   31   268.5    9.5    370  1  SEA_AVIET            P23049 avian eryth
   32   265.5    9.4    527  1  TXK_HUMAN            P42681 homo sapien
   33     261    9.2    590  1  SRC2_DROME           P08630 drosophila
   34   260.5    9.2    412  1  APKB_ARATH           P46573 arabidopsis
   35     259    9.2    901  1  CRI4_MAIZE           O24585 zea mays (m
   36     254    9.0    693  1  IRA1_MOUSE           Q62406 mus musculu
   37   251.5    8.9   2554  1  7LES_DROME           P13368 drosophila
   38     251    8.9   1370  1  IG1R_RAT             P24062 rattus norv
   39   250.5    8.9    659  1  BTK_MOUSE            P35991 mus musculu
   40     249    8.8    984  1  EPA3_RAT             O08680 rattus norv
   41   248.5    8.8    942  1  TMK1_ARATH           P43298 arabidopsis
   42   247.5    8.7    911  1  DDR1_MOUSE           Q03146 mus musculu
   43   247.5    8.7   1013  1  EPA5_CHICK           P54755 gallus gall
   44   247.5    8.7   1372  1  INSR_MOUSE           P15208 mus musculu
   45     247    8.7    980  1  CLV1_ARATH           Q9syq8 arabidopsis

                              ALIGNMENTS

RESULT   1
ANR3_HUMAN
ID   ANR3_HUMAN     STANDARD;      PRT;   832 AA.
AC   P57078;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   SERINE/THREONINE-PROTEIN KINASE ANKRD3 (EC 2.7.1.-) (ANKYRIN REPEAT
DE   DOMAIN PROTEIN 3).
GN   ANKRD3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20289799; PubMed=10830953;
RA   Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA   Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA   Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA   Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA   Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA   Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA   Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA   Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA   Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA   Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA   Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA   Lehrach H., Reinhardt R., Yaspo M.-L.;
RT   "The DNA sequence of human chromosome 21.";
RL   Nature 405:311-319(2000).
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC   -!- SIMILARITY: CONTAINS 10 ANK REPEATS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AP001743; BAA95526.1; -.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 9.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;
KW   ANK repeat.
FT   DOMAIN       22    286       PROTEIN KINASE.
FT   REPEAT      485    514       ANK 1.
FT   REPEAT      518    547       ANK 2.
FT   REPEAT      551    580       ANK 3.
FT   REPEAT      584    613       ANK 4.
FT   REPEAT      617    647       ANK 5.
```

```
FT   REPEAT      651    680       ANK 6.
FT   REPEAT      684    713       ANK 7.
FT   REPEAT      717    746       ANK 8.
FT   REPEAT      750    780       ANK 9.
FT   REPEAT      782    811       ANK 10.
FT   NP_BIND      28     36       ATP (BY SIMILARITY).
FT   BINDING      51     51       ATP (BY SIMILARITY).
FT   ACT_SITE    143    143       BY SIMILARITY.
SQ   SEQUENCE   832 AA;  91610 MW;  5D8FFFD5F04F7ECB CRC64;

  Query Match             19.1%;  Score 541.5;  DB 1;  Length 832;
  Best Local Similarity   31.7%;  Pred. No. 2.9e-30;
  Matches  171;  Conservative   77;  Mismatches  178;  Indels  113;  Gaps   23;

Qy       27 GASGTVSSARHADWRVQVAVK---HLHIHTPLLDSERKDVLREAEILHKARFSYIFPILG 83
            |  | |   ||  |:  :|:|    ||:     | || ::| ||: :' |:| || |: |
Db       31 GGFGQVYKVRHVHWKTWLAIKCSPSLHVD----DRERMELLEEAKKMEMAKFRYILPVYG 86

Qy       84 ICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLL 143
            || ||  :|:| |||  ||| :||  :     : | |||||:|| |:|:|:|| | ||||
Db       87 ICREP--VGLVMEYMETGSLEKLLASE----PLPWDLRFRIIHETAVGMNFLHCMAPPLL 140

Qy      144 HHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPEG--GTIIYMPPENYEPG 201
            | |||  |||||  :||||:||||:|   :| |   |     :|  ||| |:|||
Db      141 HLDLKPANILLDAHYHVKISDFGLAKCNGLSHSHDLSM----DGLFGTIAYLPPERIR-- 194

Qy      202 QKSRA-SIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIP 260
            :|||    |||:||:|:: | ||::|:|| |  | | ||  | :||||      ||
Db      195 EKSRLFDTKHDVYSFAIVIWGVLTQKKPFADEKNILHIMVKVVKGHRP-----ELPPVCR 249

Qy      261 HRAR----MISLIESGWAQNPDERPSF-------------LKCLIELEPVLRTFEEITFL 303
             | |    :| |::  |  :|  ||:|             |  |: :    |:  |   |
Db      250 ARPRACSHLIRLMQRCWQGDPRVRPTFQGNGLNGELIRQVLAALLPVTGRWRSPGEGFRL 309

Qy      304 EAVIQLKKT----KLQSVSSAIH-LCDK---KKMELSLNIPVNHGPQEESCGSSQLHENS 355
            |: : :: |      | ::|    ||:|   :  | : :: |   |:  |
Db      310 ESEVIIRVTCPLSSPQEITSETEDLCEKPDDEVKETAHDLDVKSPPEPRS---------E 360

Qy      356 GSPETSRSLPAPQ-DNDF-LSRKAQDCYFMKLHHCPGNHSWDSTISGSQRAAFCDHKTTP 413
              |   :   ||  |||: ||                        ||  :|     : :
Db      361 VVPARLKRASAPTFDNDYSLSELL-----------------SQLDSGVSQAVEGPEELSR 403

Qy      414 CSSAIINPLSTAGNSERLQPGIAQQWIQSKREDIVNQMTEACLNQSLDALLSRDLIMKED 473
             ||   : | ::|: :||  |::                      |:|:  |    :
Db      404 SSSE--SKLPSSGSGKRLS-GVS---------------------SVDSAFSSRGSLSLS 438

Qy      474 YELVSTKPTRTSKVRQLLDTTDIQGEEFAKVIVQKLKDNKQMGLQPYPEILVVSRSPSL 532
            :|    :|: :      | |||:| ::    ||          |||    | :    ||
Db      439 FE---REPSTSD-----LGTTDVQKKKLVDAIVSGDTSKLMKILQPQDVDLALDSGASL 489

RESULT   2
RIP_MOUSE
ID   RIP_MOUSE     .STANDARD;      PRT;   656 AA.
AC   Q60855;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   SERINE/THREONINE PROTEIN KINASE RIP (EC 2.7.1.-) (CELL DEATH PROTEIN
DE   RIP) (RECEPTOR INTERACTING PROTEIN).
GN   RIPK1 OR RIP OR RINP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6 X CBA; TISSUE=Thymus;
RX   MEDLINE=95277838; PubMed=7538908;
RA   Stanger B.Z., Leder P., Lee T.-H., Kim E., Seed B.;
RT   "RIP: a novel protein containing a death domain that interacts with
RT   Fas/APO-1 (CD95) in yeast and causes cell death.";
```

```
RL   Cell 81:513-523(1995).
CC   -!- FUNCTION: INTERACTS WITH THE DEATH DOMAIN OF FAS AND TRADD AND
CC       INITIATES APOPTOSIS. IT IS RECRUITED BY TRADD TO TNFR1 IN A TNF-
CC       DEPENDENT PROCESS. REQUIRED FOR TNFR1 ACTIVATION OF NF-KAPPA B.
CC   -!- TISSUE SPECIFICITY: FOUND AT LOW LEVELS IN ALL TISSUES.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U25995; AAB60487.1; -.
DR   HSSP; P11362; 1FGI.
DR   MGD; MGI:108212; Ripk1.
DR   InterPro; IPR000488; -.
DR   InterPro; IPR000719; -.
DR   InterPro; IPR002290; -.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding;
KW   Apoptosis.
FT   DOMAIN       17    290       PROTEIN KINASE.
FT   NP_BIND      23     31       ATP (BY SIMILARITY).
FT   BINDING      49     49       ATP (BY SIMILARITY).
FT   ACT_SITE    138    138       BY SIMILARITY.
FT   DOMAIN      568    654       DEATH DOMAIN.
FT   VARIANT     473    473       T -> I.
SQ   SEQUENCE   656 AA;  74854 MW;  ABB350B523879933 CRC64;

  Query Match             13.9%;  Score 394;  DB 1;  Length 656;
  Best Local Similarity   31.0%;  Pred. No. 4e-20;
  Matches  131;  Conservative   66;  Mismatches  144;  Indels   82;  Gaps   19;

Qy       24 LSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSERKDV-LREAEILHKARFSYIFPIL 82
            |  |  | ||   |      :  |   ::|    :|  :| | | :::|: | | :  :|
Db       23 LDSGGFGKVSLCYHRSHGFVILKK---VYTGPNRAEYNEVLLEEGKMMHRLRHSRVVKLL 79

Qy       83 GICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPL 142
            ||  |     :| |||  |:|  :|  ||:  ||   |: ||: |   |: |||:    :
Db       80 GIIIEEGNYSLVMEYMEKGNLMHVL--KTQI-DVPLSLKGRIIVEAIEGMCYLHD--KGV 134

Qy      143 LHHDLKTQNILLDNEFHVKIADFGLSKWRMMS-LSQSRSSKSAP--------EGGTIIYM 193
            :| ||| :||||:| :||:|||| |:: ::  | |:: : :|            |||: ||
Db      135 IHKDLKPENILVDRDFHIKIADLGVASFKTWSKLTKEKDNKQKEVSSTTKKNNGGTLYYM 194

Qy      194 PPE-----NYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVSQGHRP 248
             ||     | :| :||      |:||: :: | : ::|:|:| |    | :  :  |:||
Db      195 APEHLNDINAKPTEKS------DVYSFGIVLWAIFAKKEPYENVICTEQFVICIKSGNRP 248

Qy      249 VINE--ESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEP-VLRTFEEITFLEA 305
             : |  |  | :|      |||:|  |   |::||:||    |  |  |  |||    |
Db      249 NVEEILEYCPREI------ISLMERCWQAIPEDRPTFLGIEEEFRPFYLSHFEEYV-EED 301

Qy      306 VIQLKK------TKLQSVSSAIHLC---------DKKKMELSLNIPVNHGPQEESCGSSQ 350
            |  |||        || : |  | |          ::   |  :  :  || |||  ||
Db      302 VASLKKEYPDQSPVLQRMFSLQHDCVPLPPSRSNSEQPGSLHSSQGLQMGPVEESWFSS- 360

Qy      351 LHENSGSPETSRSLPAPQDNDFLSRKAQDCYFMKLHHCPGNHSWDSTISGSQRAAFCDHK 410
                  |||       ||| :  | :|      ||      |::           | : :
Db      361 ------SPE------YPQDENDRSVQA------KLQEEASYHAF---------GIFAEKQ 393

Qy      411 TTP 413
            | |
Db      394 TKP 396
```

```
RESULT   3
CTR1_ARATH
ID   CTR1_ARATH     STANDARD;      PRT;   821 AA.
AC   Q05609;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   SERINE/THREONINE-PROTEIN KINASE CTR1 (EC 2.7.1.37).
GN   CTR1 OR F17C15_150.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA; TISSUE=Seedling;
RX   MEDLINE=93161417; PubMed=8431946;
RA   Kieber J.J., Rothenberg M., Roman G., Feldmann K.A., Ecker J.R.;
RT   "CTR1, a negative regulator of the ethylene response pathway in
RT   Arabidopsis, encodes a member of the raf family of protein kinases.";
RL   Cell 72:427-441(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.-W., Rudd S.,
RA   Lemcke K., Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: ACTS AS A NEGATIVE REGULATOR IN THE ETHYLENE RESPONSE
CC       PATHWAY.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN BOTH SEEDLINGS AND ADULT PLANTS.
CC   -!- MISCELLANEOUS: CTR1 MUTANTS DISPLAY ETHYLENE-TREATED PHENOTYPES,
CC       RESULTING IN PLANTS WITH SMALL, UNEXPANDED LEAVES AND WHOSE SEED
CC       COTYLEDON GROWTH IS IMPAIRED.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       MIL/RAF SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L08789; AAA32779.1; -.
DR   EMBL; L08790; AAA32780.1; -.
DR   EMBL; AL162506; CAB82938.1; -.
DR   HSSP; P11362; 1FGI.
DR   InterPro; IPR000719; -.
DR   InterPro; IPR002290; -.
DR   Pfam; PF00069; pkinase; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding.
FT   DOMAIN       65     69       POLY-GLY.
FT   DOMAIN      135    141       POLY-GLY.
FT   DOMAIN      551    809       PROTEIN KINASE.
FT   NP_BIND     557    565       ATP (BY SIMILARITY).
FT   BINDING     578    578       ATP (BY SIMILARITY).
FT   ACT_SITE    676    676       BY SIMILARITY.
FT   MUTAGEN     596    596       E->K: IN CTR1-4; EXHIBITS ETHYLENE-
FT                                TREATED PHENOTYPE.
FT   MUTAGEN     694    694       D->E: IN CTR1-1; EXHIBITS ETHYLENE-
FT                                TREATED PHENOTYPE.
SQ   SEQUENCE   821 AA;  90306 MW;  2922D3DCD0CC15BC CRC64;

  Query Match             13.3%;  Score 376;  DB 1;  Length 821;
  Best Local Similarity   30.7%;  Pred. No. 9.6e-19;
  Matches   98;  Conservative   60;  Mismatches  111;  Indels   50;  Gaps   12;
```

```
Qy          1 MNGEAICSALPT--------------IPYHKLADLRYLSRGASGTVSSARHADWR-VQVA 45
              ||   |   :|               ||:    ||    :  :|:  :   |:|    ||
Db        520 MNAPPISQPVPNRANRELGLDGDDMDIPW---CDLNIKEKIGAGSFGTVHRAEWHGSDVA 576

Qy         46 VKHL---HIHTPLLDSER-KDVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNG 101
              || |     |     :||  : |||  |: : |   |   :|   :|  | |||||:  |
Db        577 VKILMEQDFH-----AERVNEFLREVAIMKRLRHPNIVLFMGAVTQPPNLSIVTEYLSRG 631

Qy        102 SLNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVK 161
              ||  |||:      :    |  : :::| |:|||||  ||::| |||: |:|:| :: ||
Db        632 SLYRLLHKSGAREQLDERRRLSMAYDVAKGMNYLHNRNPPIVHRDLKSPNLLVDKKYTVK 691

Qy        162 IADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPE--NYEPGQKSRASIKHDIYSYAVIT 219
              : |||||: :  :   |:|:   ||     :| ||    ||  :     | |:||: ||
Db        692 VCDFGLSRLKASTFLSSKSAAGTPE-----WMAPEVLRDEPSNE-----KSDVYSFGVIL 741

Qy        220 WEVLSRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIPH--RARMISLIESGWAQNP 277
              ||: : :||: :: || |:: :|    : :        :||     :: ::||  |   |
Db        742 WELATLQQPWGNL-NPAQVVAAVGFKCKRL--------EIPRNLNPQVAAIIEGCWTNEP 792

Qy        278 DERPSFLKCLIELEPVLRT 296
               :||||   :  | |::::
Db        793 WKRPSFATIMDLLRPLIKS 811


RESULT   4
RIP_HUMAN
ID   RIP_HUMAN      STANDARD;      PRT;   671 AA.
AC   Q13546; Q13180;
DT   01-NOV-1997 (Rel. 35, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   SERINE/THREONINE PROTEIN KINASE RIP (EC 2.7.1.-) (CELL DEATH PROTEIN
DE   RIP) (RECEPTOR INTERACTING PROTEIN).
GN   RIPK1 OR RIP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Umbilical vein endothelial cells;
RX   MEDLINE=96200892; PubMed=8612133;
RA   Hsu H., Huang J., Shu H.-B., Baichwal V.R., Goeddel D.V.;
RT   "TNF-dependent recruitment of the protein kinase RIP to the TNF
RT   receptor-1 signaling complex.";
RL   Immunity 4:387-396(1996).
RN   [2]
RP   REVISION TO 120.
RA   Huang J., Hsu H., Baichwal V.R., Goeddel D.V.;
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 300-671 FROM N.A.
RX   MEDLINE=95277838; PubMed=7538908;
RA   Stanger B.Z., Leder P., Lee T.-H., Kim E., Seed B.;
RT   "RIP: a novel protein containing a death domain that interacts with
RT   Fas/APO-1 (CD95) in yeast and causes cell death.";
RL   Cell 81:513-523(1995).
CC   -!- FUNCTION: INTERACTS WITH THE DEATH DOMAIN OF FAS AND TRADD AND
CC       INITIATES APOPTOSIS. IT IS RECRUITED BY TRADD TO TNFR1 IN A TNF-
CC       DEPENDENT PROCESS. REQUIRED FOR TNFR1 ACTIVATION OF NF-KAPPA B.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U50062; AAC32232.1; -.
```

```
DR   EMBL; U25994; AAC50137.1; -.
DR   HSSP; P11362; 1FGI.
DR   MIM; 603453; -.
DR   InterPro; IPR000488; -.
DR   InterPro; IPR000719; -.
DR   InterPro; IPR002290; -.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding;
KW   Apoptosis.
FT   DOMAIN       17    289       PROTEIN KINASE.
FT   NP_BIND      23     31       ATP (BY SIMILARITY).
FT   BINDING      49     49       ATP (BY SIMILARITY).
FT   ACT_SITE    138    138       BY SIMILARITY.
FT   DOMAIN      583    669       DEATH DOMAIN.
FT   DOMAIN      411    414       POLY-ARG.
FT   CONFLICT    514    514       T -> S (IN REF. 3).
SQ   SEQUENCE   671 AA;  75958 MW;  BADC4E7E70456ABE CRC64;

  Query Match             13.0%;  Score 369;  DB 1;  Length 671;
  Best Local Similarity   28.8%;  Pred. No. 2.2e-18;
  Matches  112;  Conservative   67;  Mismatches  146;  Indels   64;  Gaps   12;

Qy         24 LSRGASGTVSSARHADWRVQ-VAVKHLHIHTPLLDSERKDVLREAEILHKARFSYIFPIL 82
              |  |  | ||   |   | | : :       |      : :| ||::::: | | :  :|
Db         23 LDSGGFGKVSLCFH---RTQGLMIMKTVYKGPNCIEHNEALLEEAKMMNRLRHSRVVKLL 79

Qy         83 GICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPL 142
              |:  |     :| |||  |:|  :|  :   |    ::  ||: ||  |: |||     :
Db         80 GVIIEEGKYSLVMEYMEKGNLMHVLKAEMSTP---LSVKGRIILEIIEGMCYLHG--KGV 134

Qy        143 LHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPE--------GGTIIYMP 194
              :| ||| :|||:||:||:|||| ||: ::| |   :       |        |||: ||
Db        135 IHKDLKPENILVDNDFHIKIADLGLASFKMWSKLNNEEHNELREVDGTAKKNGGTLYYMA 194

Qy        195 PE-----NYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVSQGHRPV 249
              ||     | :| :||      |:||:||: | : : |:|:|:     |::  :  |:||
Db        195 PEHLNDVNAKPTEKS------DVYSFAVVLWAIFANKEPYENAICEQQLIMCIKSGNRPD 248

Qy        250 INE--ESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEITFLEAVI 307
              :::  |  | :|      |||::  |  ||: ||:|     :  |   :  | :  | |
Db        249 VDDITEYCPREI------ISLMKLCWEANPEARPTFPGIEEKFRPFYLSQLEESVEEDVK 302

Qy        308 QLKK---------TKLQSV--------SSAIHLCDKKKMELSLNIPVNHGPQEESCGSSQ 350
               |||          ::||:        ||  :   ::   |  :  :  || |||  :
Db        303 SLKKEYSNENAVVKRMQSLQLDCVAVPSSRSNSATEQPGSLHSSQGLGMGPVEESWFAPS 362

Qy        351 LHENSGSPETSRSLPAPQDNDFLSRKAQD 379
              |             |  ::   |  | ||
Db        363 LEH-----------PQEENEPSLQSKLQD 380


RESULT   5
KYK1_DICDI
ID   KYK1_DICDI     STANDARD;      PRT;  1584 AA.
AC   P18160;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   NON-RECEPTOR TYROSINE KINASE SPORE LYSIS A (EC 2.7.1.112) (TYROSINE-
DE   PROTEIN KINASE 1).
GN   PYKA OR SPLA OR DPYK1.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JH10;
RX   MEDLINE=97053827; PubMed=8898241;
```

```
RA   Nuckolls G.H., Osherov N., Loomis W.F., Spudich J.A.;
RT   "The Dictyostelium dual-specificity kinase splA is essential for
RT   spore differentiation.";
RL   Development 122:3295-3305(1996).
RN   [2]
RP   SEQUENCE OF 1248-1584 FROM N.A.
RX   MEDLINE=90287147; PubMed=1972546;
RA   Tan J.L., Spudich J.A.;
RT   "Developmentally regulated protein-tyrosine kinase genes in
RT   Dictyostelium discoideum.";
RL   Mol. Cell. Biol. 10:3578-3583(1990).
CC   -!- FUNCTION: ESSENTIAL FOR SPORE DIFFERENTIATION.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT WITH A PEAK
CC       DURING THE MOUND STAGE OF MORPHOGENESIS.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES BUT ALSO TO
CC       SERINE/THREONINE PROTEIN KINASES.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U32174; AAB41125.1; -.
DR   EMBL; M33785; AAA33202.1; -.
DR   PIR; A35670; A35670.
DR   HSSP; P11362; 1FGI.
DR   DictyDb; DD03010; pykA.
DR   InterPro; IPR000107; -.
DR   InterPro; IPR000719; -.
DR   InterPro; IPR001245; -.
DR   InterPro; IPR001660; -.
DR   Pfam; PF00536; SAM; 1.
DR   Pfam; PF00622; SPRY; 3.
DR   Pfam; PF00069; pkinase; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT   DOMAIN      403    420       POLY-ASN.
FT   DOMAIN      428    435       POLY-THR.
FT   DOMAIN      449    480       POLY-ASN.
FT   DOMAIN      483    491       POLY-ASN.
FT   DOMAIN      494    508       POLY-ASN.
FT   DOMAIN      512    532       POLY-ASN.
FT   DOMAIN      596    600       POLY-ASN.
FT   DOMAIN      808    811       POLY-PHE.
FT   DOMAIN     1026   1029       POLY-SER.
FT   DOMAIN     1195   1210       POLY-ASN.
FT   DOMAIN     1215   1220       POLY-GLN.
FT   DOMAIN     1224   1233       POLY-GLN.
FT   DOMAIN     1266   1274       POLY-PRO.
FT   DOMAIN     1289   1561       PROTEIN KINASE.
FT   NP_BIND    1295   1303       ATP (BY SIMILARITY).
FT   BINDING    1316   1316       ATP (BY SIMILARITY).
FT   ACT_SITE   1417   1417       BY SIMILARITY.
FT   CONFLICT   1248   1248       D -> R (IN REF. 2).
FT   CONFLICT   1435   1435       V -> L (IN REF. 2).
SQ   SEQUENCE   1584 AA;  174304 MW;  5D1589458D8E01E3 CRC64;

  Query Match             11.1%;  Score 313.5;  DB 1;  Length 1584;
  Best Local Similarity   28.1%;  Pred. No. 5.3e-14;
  Matches   91;  Conservative   59;  Mismatches  119;  Indels   55;  Gaps   10;

Qy         13 IPYHKLADLRYLSRGASGTVSSARHADWR-VQVAVKHLHIHTPLLDSERKDVLREAEILH 71
              | :::|   : : :|  | |   :   ||   ||:| ::       |       |  ||
Db       1284 IDFNELEFGQTIGKGFFGEV---KRGYWRETDVAIKIIYRDQFKTKSSLVMFQNEVGILS 1340
```

```
Qy     72 KARFSYIFPILGICNE--PEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILHEIA 129
          | |   :   || |     :   ||||:|  ||| : |       :    :| ::  :||
Db   1341 KLRHPNVVQFLGACTAGGEDHHCIVTEWMGGGSLRQFLTDHFNLLEQNPHIRLKLALDIA 1400

Qy    130 LGVNYLHNMTPPLLHHDLKTQNILLDN-------------EFHVKIADFGLSKWRMMSLS 176
           |:||||  |||:|| || ::|||||:             :   ||:|||||:     |
Db   1401 KGMNYLHGWTPPILHRDLSSRNILLDHNIDPKNPVVSSRQDIKCKISDFGLSR-----LK 1455

Qy    177 QSRSSKSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPL 236
          : ::|:     | | || || ::    |    | |:||| :: :|:|:  :| :|:  |:
Db   1456 KEQASQMTQSVGCIPYMAPEVFKGDSNSE---KSDVYSYGMVLFELLTSDEPQQDM-KPM 1511

Qy    237 QIMY-SVSQGHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLR 295
          :: : :  : :||       |  :   ::   ::   |  ||| ||
Db   1512 KMAHLAAYESYRP-------PIPLTTSSKWKEILTQCWDSNPDSRP-------------- 1550

Qy    296 TFEEITFLEAVIQLKKTKLQSVSS 319
               || : :: ||: : | |||
Db   1551 -----TFKQIIVHLKEMEDQGVSS 1569


RESULT   6
M3K7_MOUSE
ID   M3K7_MOUSE     STANDARD;      PRT;   579 AA.
AC   Q62073;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   MITOGEN-ACTIVATED PROTEIN KINASE KINASE KINASE 7 (EC 2.7.1.-)
DE   (TRANSFORMING GROWTH FACTOR-BETA-ACTIVATED KINASE 1) (TGF-BETA-
DE   ACTIVATED KINASE 1).
GN   MAP3K7 OR TAK1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96123277; PubMed=8533096;
RA   Yamaguchi K., Shirakabe K., Shibuya H., Irie K., Ohishi I., Ueno N.,
RA   Taniguchi T., Nishida E., Matsumoto K.;
RT   "Identification of a member of the MAPKKK family as a potential
RT   mediator of TGF-beta signal transduction.";
RL   Science 270:2008-2011(1995).
CC   -!- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKKS.
CC       MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B
CC       ACTIVATION.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       MAP KINASE KINASE KINASE SUBFAMILY.
CC   -!- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C21ORF7.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D76446; BAA11184.1; -.
DR   MGD; MGI:1346877; Map3k7.
DR   InterPro; IPR000719; -.
DR   InterPro; IPR002290; -.
DR   Pfam; PF00069; pkinase; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding.
FT   DOMAIN        8     16       POLY-SER.
FT   DOMAIN       36    291       PROTEIN KINASE.
FT   NP_BIND      42     50       ATP (BY SIMILARITY).
FT   BINDING      63     63       ATP (BY SIMILARITY).
FT   ACT_SITE    156    156       BY SIMILARITY.
```

```
SQ   SEQUENCE   579 AA;  64227 MW;  97C8F6F3C8E283EE CRC64;

  Query Match             11.0%;  Score 311;  DB 1;  Length 579;
  Best Local Similarity   23.5%;  Pred. No. 2e-14;
  Matches  136;  Conservative   97;  Mismatches  246;  Indels  100;  Gaps   22;

Qy     13 IPYHKLADLRYLSRGASGTVSSARHADWRVQ-VAVKHLHIHTPLLDSERKDVLREAEILH 71
          | | ::     : ||| | |  |:   || : ||:| :       :||||  : |   |
Db     31 IDYKEIEVEEVVGRGAFGVVCKAK---WRAKDVAIKQIE-----SESERKAFIVELRQLS 82

Qy     72 KARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYP------DVAWPLRFRIL 125
          :     |  : | |  |  : :| ||   |||  :||     |       ::| |
Db     83 RVNHPNIVKLYGACLNP--VCLVMEYAEGGSLYNVLHGAEPLPYYTAAHAMSWCL----- 135

Qy    126 HEIALGVNYLHNMTP-PLLHHDLKTQNILL-DNEFHVKIADFGLSKWRMMSLSQSRSSKS 183
           : : || |||:| |  |:| |||  |:||      :|| ||| :      :: ::
Db    136 -QCSQGVAYLHSMQPKALIHRDLKPPNLLLVAGGTVLKICDFGTACDIQTHMTNNK---- 190

Qy    184 APEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNP-LQIMYSV 242
              |:  :| || :|    |  | | |::|: :| |||::|::||:::  |  :||::|
Db    191 ----GSAAWMAPEVFE---GSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWAV 243

Qy    243 SQGHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTF----E 298
            | || :  ::||  |       ||:   |:::| :|||  : :  :  ::| |    |
Db    244 HNGTRPPL-IKNLPKPIE------SLMTRCWSKDPSQRPSMEEIVKIMTHLMRYFPGADE 296

Qy    299 EI-------------------TFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLNIPVNH 339
           :                   :|::        |  :    :   :     |   :  |
Db    297 PLQYPCQYSDEGQSNSATSTGSFMDIASTNTSNKSDTNMEQVPATNDTIKRLESKLLKNQ 356

Qy    340 GPQEESCGSSQLHENSGSPETSRSLPAPQDNDFLSRKAQDCYFMKLHHCPGNHS------ 393
            |:   |   |  : ||  :  |||   :   :|    :    ::    ||
Db    357 AKQQSESGRLSLGASRGS--SVESLPPTSEGKRMSADMSEIE-ARIVATAGNGQPRRRSI 413

Qy    394 WDSTISGSQRAAFCDHKTTPCSSAIINPLSTAGNSERLQPGIAQQWIQSKREDIVNQMTE 453
           | |::|::        ::|    |     |:    |  |        :   |    |
Db    414 QDLTVTGTEPGQVSSRSSSPSVRMITTSGPTSEKPARSHPWTPDDSTDTNGSDNSIPMAY 473

Qy    454 ACLNQSLD------------ALLSRDLIMKEDYELVSTKPTRTSKVRQLLDTTDIQGEEF 501
            |:  |             |:  :   | ::|  | |      ::  ||     : :|
Db    474 LTLDHQLQPLAPCPNSKESMAVFEQHCKMAQEYMKVQT------EIALLLQ----RKQEL 523

Qy    502 AKVIVQKLKD--NKQMGLQPYPEILVVSRSPSLNLLQNK 538
             : |  ||  |    :| : ::|  ::| |    | |
Db    524 VAELDQDEKDQQNTSRLVQEHKKLLDENKSLSTYYQQCK 562


RESULT   7
KYK2_DICDI
ID   KYK2_DICDI     STANDARD;      PRT;   410 AA.
AC   P18161;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   TYROSINE-PROTEIN KINASE 2 (EC 2.7.1.112) (FRAGMENT).
GN   PYKB OR DPYK2.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90287147; PubMed=1972546;
RA   Tan J.L., Spudich J.A.;
RT   "Developmentally regulated protein-tyrosine kinase genes in
RT   Dictyostelium discoideum.";
RL   Mol. Cell. Biol. 10:3578-3583(1990).
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES BUT ALSO TO
CC       SERINE/THREONINE PROTEIN KINASES.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M33784; AAA33203.1; -.
DR   PIR; B35670; B35670.
DR   HSSP; P11362; 1FGI.
DR   DictyDb; DD03011; pykB.
DR   InterPro; IPR000719; -.
DR   InterPro; IPR001245; -.
DR   Pfam; PF00069; pkinase; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT   NON_TER       1      1
FT   DOMAIN      108    381       PROTEIN KINASE.
FT   NP_BIND     114    122       ATP (BY SIMILARITY).
FT   BINDING     135    135       ATP (BY SIMILARITY).
FT   ACT_SITE    232    232       BY SIMILARITY.
SQ   SEQUENCE   410 AA;  46386 MW;  E93918B605B9AEC1 CRC64;

  Query Match             10.7%;  Score 303.5;  DB 1;  Length 410;
  Best Local Similarity   30.1%;  Pred. No. 4.2e-14;
  Matches   87;  Conservative   57;  Mismatches  102;  Indels   43;  Gaps   13;

Qy       20 DLRYLSRGASGTVSSARHADWR-VQVAVKHLHIHTPLLDSE--RKDVLREAEILHKARFS 76
            |:::: :   |  |      |: : ||:| | |   : | |  ::  :|| : | |
Db      107 DIQFIQKVGEGAFSEVWEGWWKGIHVAIKKLKI---IGDEEQFKERFIREVQNLKKGNHQ 163

Qy       77 YIFPILGICNEPEFLGIVTEYMPNGSLNELLHR------KTEYPDVAWPLRFRILHEIAL 130
             |   :| | :|    |:||||  |||  :||       | :|   :||  ::  ::||
Db      164 NIVMFIGACYKPAC--IITEYMAGGSLYNILHNPNSSTPKVKY---SFPLVLKMATDMAL 218

Qy      131 GVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPEGGTI 190
            |: :||::|  ::| || :||||||   ::||:||||      |  :||        | |
Db      219 GLLHLHSIT--IVHRDLTSQNILLDELGNIKISDFGL------SAEKSREGSMTMTNGGI 270

Qy      191 I---YMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDV---TNPLQIMYSVSQ 244
                : |||  :       | | |:| :::: ||:|: : || |:       |: |:
Db      271 CNPRWRPPELTK--NLGHYSEKVDVYCFSLVVWEILTGEIPFSDLDGSQRSAQVAYA--- 325

Qy      245 GHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPV 293
            | || | |   |        :  |:   |  :|::|| |   : :|: :
Db      326 GLRPPIPEYCDP-------ELKLLLTQCWEADPNDRPPFTYIVNKLKEI 367

RESULT   8
M3K7_HUMAN
ID   M3K7_HUMAN     STANDARD;      PRT;   606 AA.
AC   O43318; O43317; O43319;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   MITOGEN-ACTIVATED PROTEIN KINASE KINASE KINASE 7 (EC 2.7.1.-)
DE   (TRANSFORMING GROWTH FACTOR-BETA-ACTIVATED KINASE 1) (TGF-BETA-
DE   ACTIVATED KINASE 1).
GN   MAP3K7 OR TAK1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC   TISSUE=Lung;
RX   MEDLINE=98153801; PubMed=9480845;
RA   Sakurai H., Shigemori N., Hasegawa K., Sugita T.;
RT   "TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an
RT   NF-kappa B-inducing kinase-independent mechanism.";
RL   Biochem. Biophys. Res. Commun. 243:545-549(1998).
CC   -!- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKKS.
CC       MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B
CC       ACTIVATION.
CC   -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1A, 1B (SHOWN HERE) AND 1C; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       MAP KINASE KINASE KINASE SUBFAMILY.
CC   -!- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C21ORF7.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AB009357; BAA25026.1; -.
DR   EMBL; AB009356; BAA25025.1; -.
DR   EMBL; AB009358; BAA25027.2; -.
DR   MIM; 602614; -.
DR   InterPro; IPR000719; -.
DR   InterPro; IPR002290; -.
DR   Pfam; PF00069; pkinase; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding;
KW   Alternative splicing.
FT   DOMAIN        8     14       POLY-SER.
FT   DOMAIN       36    291       PROTEIN KINASE.
FT   NP_BIND      42     50       ATP (BY SIMILARITY).
FT   BINDING      63     63       ATP (BY SIMILARITY).
FT   ACT_SITE    156    156       BY SIMILARITY.
FT   VARSPLIC    404    430       MISSING (IN ISOFORM 1A).
FT   VARSPLIC    509    518       PLAPCPNSKE -> ARTSCRTGPG (IN ISOFORM 1C).
FT   VARSPLIC    519    606       MISSING (IN ISOFORM 1C).
SQ   SEQUENCE   606 AA;  67196 MW;  3D8F8147CD174013 CRC64;

  Query Match             10.6%;  Score 300.5;  DB 1;  Length 606;
  Best Local Similarity   22.8%;  Pred. No. 1.2e-13;
  Matches  138;  Conservative   96;  Mismatches  246;  Indels  125;  Gaps   23;

Qy       13 IPYHKLADLRYLSRGASGTVSSARHADWRVQ-VAVKHLHIHTPLLDSERKDVLREAEILH 71
            | | ::     : ||| | |  |:   || : ||:| :       :||||  : |   |
Db       31 IDYKEIEVEEVVGRGAFGVVCKAK---WRAKDVAIKQIE-----SESERKAFIVELRQLS 82

Qy       72 KARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYP------DVAWPLRFRIL 125
            :     |  : | |  |  : :| ||   |||  :||     |       ::| |
Db       83 RVNHPNIVKLYGACLNP--VCLVMEYAEGGSLYNVLHGAEPLPYYTAAHAMSWCL----- 135

Qy      126 HEIALGVNYLHNMTP-PLLHHDLKTQNILL-DNEFHVKIADFGLSKWRMMSLSQSRSSKS 183
             : : || |||:| |  |:| |||  |:||      :|| ||| :      :: ::
Db      136 -QCSQGVAYLHSMQPKALIHRDLKPPNLLLVAGGTVLKICDFGTACDIQTHMTNNK---- 190

Qy      184 APEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNP-LQIMYSV 242
                |:  :| || :|    |  | | |::|: :| |||::|::||:::  |  :||::|
Db      191 ----GSAAWMAPEVFE---GSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWAV 243

Qy      243 SQGHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTF----E 298
              | || :  ::||  |       ||:   |:::| :|||  : :  :  ::| |    |
Db      244 HNGTRPPL-IKNLPKPIE------SLMTRCWSKDPSQRPSMEEIVKIMTHLMRYFPGADE 296

Qy      299 EI-------------------TFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLNIPVNH 339
             :                   :|::        |  :    :   :     |   :  |
Db      297 PLQYPCQYSDEGQSNSATSTGSFMDIASTNTSNKSDTNMEQVPATNDTIKRLESKLLKNQ 356

Qy      340 GPQEESCGSSQLHENSGSPETSRSLPAPQDNDFLSRKAQD-------------------- 379
              |:   |   |  : ||  :  |||   :   :|    :
Db      357 AKQQSESGRLSLGASRGS--SVESLPPTSEGKRMSADMSEIEARIAATTAYSKPKRGHRK 414
```

```
Qy      380 -CYFMKLHHCP-----GNHS------WDSTISGSQRAAFCDHKTTPCSSAIINPLSTAGN 427
               |  :   |     ||         | |::|::        ::|    |     |:
Db      415 TASFGNILDVPEIVISGNGQPRRRSIQDLTVTGTEPGQVSSRSSSPSVRMITTSGPTSEK 474

Qy      428 SERLQPGIAQQWIQSKREDIVNQMTEACLNQSLD------------ALLSRDLIMKEDYE 475
              |  |        :   |    |    |:  |             |:  :   | ::|
Db      475 PTRSHPWTPDDSTDTNGSDNSIPMAYLTLDHQLQPLAPCPNSKESMAVFEQHCKMAQEYM 534

Qy      476 LVSTKPTRTSKVRQLLDTTDIQGEEFAKVIVQKLKD--NKQMGLQPYPEILVVSRSPSLN 533
             | |      ::  ||     : :|    : |  ||  |    :|  : ::|  ::| |
Db      535 KVQT------EIALLLQ----RKQELVAELDQDEKDQQNTSRLVQEHKKLLDENKSLSTY 584

Qy      534 LLQNK 538
              | |
Db      585 YQQCK 589

RESULT   9
ANPB_ANGJA
ID   ANPB_ANGJA     STANDARD;      PRT;  1050 AA.
AC   P55202;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   ATRIAL NATRIURETIC PEPTIDE RECEPTOR B PRECURSOR (ANP-B) (ANPRB) (GC-B)
DE   (GUANYLATE CYCLASE) (EC 4.6.1.2) (NPR-B) (ATRIAL NATRIURETIC PEPTIDE
DE   B-TYPE RECEPTOR).
OS   Anguilla japonica (Japanese eel).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguilloidei;
OC   Anguillidae; Anguilla.
OX   NCBI_TaxID=7937;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Gill;
RX   MEDLINE=94298823; PubMed=7913035;
RA   Katafuchi T., Takashima A., Kashiwagi M., Hagiwara H., Takei Y.,
RA   Hirose S.;
RT   "Cloning and expression of eel natriuretic-peptide receptor B and
RT   comparison with its mammalian counterparts.";
RL   Eur. J. Biochem. 222:835-842(1994).
CC   -!- FUNCTION: RECEPTOR FOR NATRIURETIC PEPTIDES. HAS GUANYLATE
CC       CYCLASE ACTIVITY ON BINDING TO LIGAND. THE ORDER OF POTENCY OF
CC       LIGANDS IN STIMULATING GC ACTIVITY IS CNP > VNP > ANP.
CC   -!- CATALYTIC ACTIVITY: GTP = 3,'5'-CYCLIC GMP + PYROPHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: HIGH LEVELS FOUND IN LIVER, ATRIUM AND GILL.
CC       MODERATE LEVELS FOUND IN BRAIN AND VENTRICLE, AND LOW LEVELS IN
CC       OESOPHAGEAL SPHINCTER, STOMACH, POSTERIOR INTESTINE AND KIDNEY.
CC   -!- INDUCTION: BY OSMOSIS. LEVELS DECREASE UNDER SALINE CONDITIONS.
CC   -!- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO
CC       WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C)
CC       WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE
CC       CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.
CC   -!- SIMILARITY: SOME SIMILARITY WITH CONSERVED REGION OF CATALYTIC
CC       DOMAIN OF PROTEIN KINASES.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   -!- SIMILARITY: TO ANP-A ON THE COMPLETE SEQUENCE, AND TO ANP-C IN
CC       THE EXTRACELLULAR AND TRANSMEMBRANE DOMAINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D25417; BAA05007.1; -.
DR   HSSP; Q02846; 1AWL.
DR   InterPro; IPR000719; -.
DR   InterPro; IPR001054; -.
```

```
DR   InterPro; IPR001170; -.
DR   InterPro; IPR001828; -.
DR   Pfam; PF01094; ANF_receptor; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR00255; NATPEPTIDER.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;
KW   cGMP synthesis; Signal.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20   1050       ATRIAL NATRIURETIC PEPTIDE RECEPTOR B.
FT   DOMAIN       20    460       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    461    481       POTENTIAL.
FT   DOMAIN      482   1050       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      517    790       PROTEIN KINASE LIKE.
FT   DISULFID     84    110       BY SIMILARITY.
FT   DISULFID    236    339       BY SIMILARITY.
FT   DISULFID    443    443       INTERCHAIN (PROBABLE).
FT   DISULFID    452    452       INTERCHAIN (PROBABLE).
FT   CARBOHYD     26     26       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     74     74       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    169    169       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    203    203       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    352    352       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    366    366       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    415    415       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1050 AA;  119857 MW;  F3AC6DDD17BD3832 CRC64;

  Query Match             10.4%;  Score 295;  DB 1;  Length 1050;
  Best Local Similarity   29.0%;  Pred. No. 5.9e-13;
  Matches   87;  Conservative   65;  Mismatches   90;  Indels   58;  Gaps   16;

Qy       15 YHKLADLRYL--SRGAS-GTVSSARHADWRV----------QVAVKHLHIHTPLLDSERK 61
            ||| |  |     ||:| |:: :| |  :::           ||:|  |::   ::  |:
Db      511 YHKCAGSRLTISQRGSSYGSLITA-HGKYQLFAKTGYFKGNLVAIK--HVNKKRIELTRQ 567

Qy       62 DVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLR 121
             || | : :   :|:::   :| | :|  : ||||| | ||| ::|  : |  :: |  |
Db      568 -VLFELKHMRDVQFNHLTRFIGACIDPPNICIVTEYCPRGSLQDIL--ENESINLDWMFR 624

Qy      122 FRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQ---- 177
            : ::::|  |:|:|||      | :||: | ::|: | :|| |:||: :|    ::
Db      625 YSLINDIVKGMNFLHNSYIG-SHGNLKSSNCVVDSRFVLKITDYGLASFRSSCENEDSHA 683

Qy      178 --SRSSKSAPEGGTIIY--MPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPF---- 229
              ::   :|||   :||   ||:  : |         |:||: :|  |:  |  ||
Db      684 LYAKKLWTAPE--LLIYDRHPPQGTQKG---------DVYSFGIILQEIALRNGPFYVDG 732

Qy      230 EDVTNPLQIMYSVSQGHRPVINEESLPYDIP------HRARMISLIESGWAQNPDERPSF 283
             |:  :| :|:  |  | :        ||  |      |   :  |:|  ||::| :|| |
Db      733 MDL-SPKEIVQKVRNGQK--------PYFRPTTDTSCHSEELSILMEGCWAEDPADRPDF 783

RESULT  10
ANPB_RAT
ID   ANPB_RAT       STANDARD;      PRT;  1047 AA.
AC   P16067;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   ATRIAL NATRIURETIC PEPTIDE RECEPTOR B PRECURSOR (ANP-B) (ANPRB) (GC-B)
DE   (GUANYLATE CYCLASE) (EC 4.6.1.2) (NPR-B) (ATRIAL NATRIURETIC PEPTIDE
DE   B-TYPE RECEPTOR).
GN   NPR2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RX   MEDLINE=89376566; PubMed=2570641;
RA   Schulz S., Singh S., Bellet R.A., Singh G., Tubb D.J., Chin H.,
RA   Garbers D.L.;
RT   "The primary structure of a plasma membrane guanylate cyclase
RT   demonstrates diversity within this new receptor family.";
RL   Cell 58:1155-1162(1989).
CC   -!- FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. HAS GUANYLATE
CC       CYCLASE ACTIVITY ON BINDING OF ANF. SEEMS TO BE STIMULATED MORE
CC       EFFECTIVELY BY BRAIN NATRIURETIC PEPTIDE (BNP) THAN BY ANP.
CC   -!- CATALYTIC ACTIVITY: GTP = 3,'5'-CYCLIC GMP + PYROPHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO
CC       WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C)
CC       WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE
CC       CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.
CC   -!- SIMILARITY: SOME SIMILARITY WITH CONSERVED REGION OF CATALYTIC
CC       DOMAIN OF PROTEIN KINASES.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   -!- SIMILARITY: TO ANP-A ON THE COMPLETE SEQUENCE, AND TO ANP-C IN TH
CC       EXTRACELLULAR AND TRANSMEMBRANE DOMAINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M26896; AAA41205.1; -.
DR   PIR; A33300; OYRTBR.
DR   HSSP; Q02846; 1AWL.
DR   InterPro; IPR000719; -.
DR   InterPro; IPR001054; -.
DR   InterPro; IPR001170; -.
DR   InterPro; IPR001828; -.
DR   Pfam; PF01094; ANF_receptor; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR00255; NATPEPTIDER.
DR   PROSITE; PS00458; ANF_RECEPTORS; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;
KW   cGMP synthesis; Signal.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23   1047       ATRIAL NATRIURETIC PEPTIDE RECEPTOR B.
FT   DOMAIN       23    458       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    459    478       POTENTIAL.
FT   DOMAIN      479   1047       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      513    786       PROTEIN KINASE LIKE.
FT   DISULFID     75    101       BY SIMILARITY.
FT   DISULFID    439    439       INTERCHAIN (PROBABLE).
FT   DISULFID    448    448       INTERCHAIN (PROBABLE).
FT   CARBOHYD     24     24       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     35     35       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    161    161       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    195    195       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    244    244       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    277    277       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    349    349       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1047 AA;  117126 MW;  5062C49228CC14A3 CRC64;

  Query Match             10.3%;  Score 292.5;  DB 1;  Length 1047;
  Best Local Similarity   27.3%;  Pred. No. 8.8e-13;
  Matches  102;  Conservative   80;  Mismatches  114;  Indels   77;  Gaps   21;

Qy         15 YHKLADLRY-LS-RGAS-GTVSSARHADWRV----------QVAVKHLHIHTPLLDSERK 61
              ||| |  |  || ||:| |:: :| |  :::           ||:|  |::   ::  |:
Db        507 YHKGAGSRLTLSLRGSSYGSLMTA-HGKYQIFANTGHFKGNVVAIK--HVNKKRIELTRQ 563

Qy         62 DVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLR 121
               || | : :   :|:::   :| | :|  : ||||| | ||| ::|  : :  :: |  |
Db        564 -VLFELKHMRDVQFNHLTRFIGACIDPPNICIVTEYCPRGSLQDIL--ENDSINLDWMFR 620

Qy        122 FRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSS 181
              : :::::  |: :|||      |  ||: | ::|: | :|| |:||: :|         |
Db        621 YSLINDLVKGMAFLHNSIIS-SHGSLKSSNCVVDSRFVLKITDYGLASFR---------S 670

Qy        182 KSAPEGGTIIY-----MPPE----NYEPGQKSRASIKHDIYSYAVITWEVLSRKQPF--E 230
               : |:    :|       ||    |  |   :    | |:||:|:|  |:  |  ||  |
Db        671 TAEPDDSHALYAKKLWTAPELLSGNPLP---TTGMQKADVYSFAIILQEIALRSGPFYLE 727

Qy        231 DV-TNPLQIMYSVSQGHRPV---------INEESLPYDIPHRARMISLIESGWAQNPDER 280
               :  :| :|:  |  | ||          :|||           :: |:|  |||:| ||
Db        728 GLDLSPKEIVQKVRNGQRPYFRPSIDRTQLNEE-----------LVLLMERCWAQDPTER 776

Qy        281 PSFLKCLIELEPVLRTFEE---ITFLE-AVIQLKK-----TKLQSVSSAIHLCDKKKMEL 331
              | |     :::  :| | :    : |:  ::::::      ||    :  :| :|:| |
Db        777 PDF----GQIKGFIRRFNKEGGTSILDNLLLRMEQYANNLEKLVEERTQAYLEEKRKAEA 832

Qy        332 SLNIPVNHGPQEE 344
               |   : |   |:
Db        833 LLYQILPHSVAEQ 845


RESULT  11
M3KA_HUMAN
ID   M3KA_HUMAN      STANDARD;      PRT;   954 AA.
AC   Q02779; Q12761; Q14871;
DT   01-FEB-1994 (Rel. 28, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   MITOGEN-ACTIVATED PROTEIN KINASE KINASE KINASE 10 (EC 2.7.1.-)
DE   (MIXED LINEAGE KINASE 2) (PROTEIN KINASE MST).
GN   MAP3K10 OR MLK2 OR MST.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=96128179; PubMed=8536694;
RA   Dorow D.S., Devereux L., Tu G.F., Price G., Nicholl J.K.,
RA   Sutherland G.R., Simpson R.J.;
RT   "Complete nucleotide sequence, expression, and chromosomal
RT   localisation of human mixed-lineage kinase 2.";
RL   Eur. J. Biochem. 234:492-500(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=95249256; PubMed=7731697;
RA   Katoh M., Hirai M., Sugimura T., Terada M.;
RT   "Cloning and characterization of MST, a novel (putative)
RT   serine/threonine kinase with SH3 domain.";
RL   Oncogene 10:1447-1451(1995).
RN   [3]
RP   SEQUENCE OF 244-480 FROM N.A.
RC   TISSUE=Colon epithelium;
RX   MEDLINE=93238756; PubMed=8477742;
RA   Dorow D.S., Devereux L., Dietzsch E., de Kretser T.;
RT   "Identification of a new family of human epithelial protein kinases
RT   containing two leucine/isoleucine-zipper domains.";
RL   Eur. J. Biochem. 213:701-710(1993).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND SKELETAL MUSCLE.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       MAP KINASE KINASE KINASE SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; X90846; CAA62351.1; -.
DR   EMBL; Z48615; CAA88531.1; -.
DR   PIR; S32468; S32468.
DR   HSSP; P00523; 2PTK.
DR   MIM; 600137; -.
DR   InterPro; IPR000719; -.
DR   InterPro; IPR001245; -.
DR   InterPro; IPR001452; -.
DR   InterPro; IPR002965; -.
DR   Pfam; PF00018; SH3; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PRINTS; PR00452; SH3DOMAIN.
DR   PRINTS; PR01217; PRICHEXTENSN.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS50002; SH3; 1.
KW   Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
KW   ATP-binding; SH3 domain.
FT   DOMAIN        2      5       POLY-GLU.
FT   DOMAIN       16     81       SH3.
FT   DOMAIN       98    360       PROTEIN KINASE.
FT   NP_BIND     104    112       ATP (BY SIMILARITY).
FT   BINDING     125    125       ATP (BY SIMILARITY).
FT   ACT_SITE    222    222       BY SIMILARITY.
FT   DOMAIN      384    405       LEUCINE-ZIPPER (BY SIMILARITY).
FT   DOMAIN      419    440       LEUCINE-ZIPPER (BY SIMILARITY).
FT   DOMAIN      449    463       ARG/LYS-RICH (BASIC).
FT   CONFLICT    462    464       SRL -> AV (IN REF. 2).
FT   CONFLICT    471    471       G -> S (IN REF. 2).
FT   CONFLICT    807    807       G -> R (IN REF. 2).
FT   CONFLICT    818    818       V -> A (IN REF. 2).
FT   CONFLICT    465    480       LKLREGGSHISLPSGF -> AQAAGRRQPHQPALWL (IN
FT                                REF. 3).
SQ   SEQUENCE   954 AA;  103623 MW;  538F4AAA559B0ABA CRC64;

  Query Match             10.3%;  Score 290;  DB 1;  Length 954;
  Best Local Similarity   23.7%;  Pred. No. 1.2e-12;
  Matches  123;  Conservative   78;  Mismatches  174;  Indels  144;  Gaps   23;

Qy       13 IPYHKLADLRYLSRGASGTVSSARHADWR-VQVAVKHLHIHTPLLDSER------KDVLR 65
            ||:|:|     :  |  | |  |    ||  :||||        || |:      : | :
Db       93 IPFHELQLEEIIGVGGFGKVYRAL---WRGEEVAVK-----AARLDPEKDPAVTAEQVCQ 144

Qy       66 EAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVA--WPLRFR 123
            || :    :   |  : | |  |  | :| ||   |:|: :|  :   | |   | :
Db      145 EARLFGALQHPNIIALRGACLNPPHLCLVMEYARGGALSRVLAGRRVPPHVLVNWAV--- 201

Qy      124 ILHEIALGVNYLHNMTP-PLLHHDLKTQNILLDN--EFH------VKIADFGLSK-WRMM 173
               ::| |:|||||  | |::| |||: |||:    | |      :|| ||||:: |
Db      202 ---QVARGMNYLHNDAPVPIIHRDLKSINILILEAIENHNLADTVLKITDFGLAREWH-- 256

Qy      174 SLSQSRSSKSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVT 233
                 :::| :   ||  :| ||       |::|   |::|: |: ||:|: : |: ::
Db      257 -----KTTKMS-AAGTYAWMAPEVIRLSLFSKSS---DVWSFGVLLWELLTGEVPYREI- 306

Qy      234 NPLQIMYSVSQGHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPV 293
            : | : | |:      :|: :||           |:|  |  :|  || |   |  || :
Db      307 DALAVAYGVA------MNKLTLPIPSTCPEPFARLLEECWDPDPHGRPDFGSILKRLEVI 360

Qy      294 ---------------------------------------------LRTFEEITFLEAVI- 307
                                                         ||  :|  | |  :
Db      361 EQSALFQMPLESFHSLQEDWKLEIQHMFDDLRTKEKELRSREEELLRAAQEQRFQEEQLR 420

Qy      308 ----QLKKTKLQSVSSAIHLCDKKKMELSLNIPVNHGPQEESCGSSQLHENSGSPETSRS 363
                :| : ::  |   :||      :||   |             :: :  |: : ||
Db      421 RREQELAEREMDIVERELHLL---MCQLSQEKP-------------RVRKRKGNFKRSRL 464

Qy      364 LPAPQDNDFLSRKAQDCYFMKLHHCPGNHSWDSTISGS---QRAAFCDHKTTPCSSAII- 419
            |   :    :|              |       |:  |    :    |  : | | :||
Db      465 LKLREGGSHIS-------------LPSGFEHKITVQASPTLDKRKGSDGASPPASPSIIP 511

Qy      420 -------NPLSTAGNSERLQPGIAQQWIQS---KREDIV 448
                    |:   |:|     | :  | :    |:|::|
Db      512 RLRAIRLTPVDCGGSSSGSSSGGSGTWSRGGPPKKEELV 550


RESULT  12
RLK5_ARATH
ID   RLK5_ARATH      STANDARD;      PRT;   999 AA.
AC   P47735;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   RECEPTOR-LIKE PROTEIN KINASE 5 PRECURSOR (EC 2.7.1.-).
GN   RLK5 OR AT4G28490 OR F21O9.180.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RX   MEDLINE=94035150; PubMed=8220453;
RA   Walker J.C.;
RT   "Receptor-like protein kinase genes of Arabidopsis thaliana.";
RL   Plant J. 3:451-456(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RX   MEDLINE=20083488; PubMed=10617198;
RA   Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA   Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA   Harris B., Ansorge W., Brandt P., Grivell L., Rieger M.,
RA   Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA   Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,
RA   Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA   Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA   Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA   Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA   Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA   Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA   Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA   Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA   Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA   De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
RA   Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA   Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
RA   Pettett A., Rajandream M.-A., Lyne M., Benes V., Rechmann S.,
RA   Borkova D., Bloecker H., Scharfe M., Grimm M., Loehnert T.-H.,
RA   Dose S., de Haan M., Maarse A., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA   Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
RA   Massenet O., Quigley F., Clabauld G., Muendlein A., Felber R.,
RA   Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA   Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA   Gibbons T., Weber N., Vandenbol M., Bargues M., Terol J., Torres A.,
RA   Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA   Heijnen L., Schwarz S., Scholler P., Heber S., Francs P., Bielke C.,
RA   Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA   Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA   Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA   Sehkon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA   Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA   Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA   Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA   Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA   Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA   Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshu C.,
```

```
RA   Antonoiu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA   Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA   Swaby I.K., O'Shaugnessy A., Rodriguez M., Hoffman J., Till S.,
RA   Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA   Chen E., Marra M., Martienssen R., McCombie W.R.;
RT   "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:769-777(1999).
RN   [3]
RP   CHARACTERIZATION.
RX   MEDLINE=94368830; PubMed=8086440;
RA   Horn M.A., Walker J.C.;
RT   "Biochemical properties of the autophosphorylation of RLK5, a
RT   receptor-like protein kinase from Arabidopsis thaliana.";
RL   Biochim. Biophys. Acta 1208:65-74(1994).
CC   -!- COFACTOR: HAVE SIGNIFICANTLY GREATER ACTIVITY IN THE PRESENCE OF
CC       MN2+ THAN MG2+.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ROOTS AND ROSETTES.
CC   -!- PTM: AUTOPHOSPHORYLATED ON SERINE AND THREONINE RESIDUES.
CC   -!- SIMILARITY: TO PROTEIN KINASES IN THE C-TERMINUS. BUT DOES NOT
CC       SEEM TO HAVE CONSERVED A KINASE ACTIVITY.
CC   -!- SIMILARITY: CONTAINS 18 LEUCINE-RICH REPEATS (LRR).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M84660; AAA32859.1; -.
DR   EMBL; AL021749; CAA16889.1; -.
DR   EMBL; AL161572; CAB79651.1; -.
DR   HSSP; P00523; 2PTK.
DR   InterPro; IPR000719; -.
DR   InterPro; IPR001611; -.
DR   InterPro; IPR002290; -.
DR   Pfam; PF00560; LRR; 14.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR00019; LEURICHRPT.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding;
KW   Transmembrane; Glycoprotein; Phosphorylation; Leucine-rich repeat;
KW   Repeat; Signal.
FT   SIGNAL        1     14       POTENTIAL.
FT   CHAIN        15    999       RECEPTOR-LIKE PROTEIN KINASE 5.
FT   DOMAIN       15    621       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    622    641       POTENTIAL.
FT   DOMAIN      642    999       CYTOPLASMIC (POTENTIAL).
FT   REPEAT       88    112       LRR 1.
FT   REPEAT      114    138       LRR 2.
FT   REPEAT      138    161       LRR 3.
FT   REPEAT      163    186       LRR 4.
FT   REPEAT      188    211       LRR 5.
FT   REPEAT      235    261       LRR 6.
FT   REPEAT      263    283       LRR 7.
FT   REPEAT      283    306       LRR 8.
FT   REPEAT      306    330       LRR 9.
FT   REPEAT      332    353       LRR 10.
FT   REPEAT      354    378       LRR 11.
FT   REPEAT      402    426       LRR 12.
FT   REPEAT      426    450       LRR 13.
FT   REPEAT      452    474       LRR 14.
FT   REPEAT      498    522       LRR 15.
FT   REPEAT      524    547       LRR 16.
FT   REPEAT      549    567       LRR 17.
FT   REPEAT      567    592       LRR 18.
FT   DOMAIN      683    968       PROTEIN KINASE.
FT   CARBOHYD     98     98       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    102    102       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    150    150       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    185    185       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    210    210       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    269    269       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    282    282       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    452    452       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    576    576       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   NP_BIND     689    697       ATP (BY SIMILARITY).
FT   BINDING     711    711       ATP (BY SIMILARITY).
FT   ACT_SITE    819    819       BY SIMILARITY.
FT   MUTAGEN     711    711       K->E: LOSS OF CATALYTIC ACTIVITY.
SQ   SEQUENCE   999 AA;  109095 MW;  F5793D899EA0C6A7 CRC64;

  Query Match             10.2%;  Score 289.5;  DB 1;  Length 999;
  Best Local Similarity   32.3%;  Pred. No. 1.3e-12;
  Matches   97;  Conservative   49;  Mismatches  123;  Indels   31;  Gaps   12;

Qy     16 HKLADL----RYLSRGASGTVSSARHADWRVQVAVKHLHIHTP------LLDSERKDVL- 64
          |::||       :  |:|| |         | |||| |:            ||  :||
Db    677 HEIADCLDEKNVIGFGSSGKVYKVELRGGEV-VAVKKLNKSVKGGDDEYSSDSLNRDVFA 735

Qy     65 REAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLH--RKTEYPDVAWPLRF 122
           | | |   |   |  :   |:  :   :| |||||||| ::||  ||     : || |
Db    736 AEVETLGTIRHKSIVRLWCCCSSGDCKLLVYEYMPNGSLADVLHGDRKGGVV-LGWPERL 794

Qy    123 RILHEIALGVNYL-HNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSS 181
          ||  : | |::|| |:  ||::| |:|: |||||:::  |:||||::|   ||  |::  :
Db    795 RIALDAAEGLSYLHHDCVPPIVHRDVKSSNILLDSDYGAKVADFGIAKVGQMSGSKTPEA 854

Qy    182 KSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYS 241
           |    |:  |: ||        | : | ||||: |:  |::: ||| :       :
Db    855 MSG-IAGSCGYIAPEYV---YTLRVNEKSDIYSFGVVLLELVTGKQPTDSELGDKDMAKW 910

Qy    242 VSQ-----GHRPVINEESLPYDIPHRARMISLIESG---WAQNPDERPSFLKCLIELEPV 293
          |       |  |||: :    |:  :  :  :|  |    :  |  |||  | :| |: |
Db    911 VCTALDKCGLEPVIDPK---LDLKFKEEISKVIHIGLLCTSPLPLNRPSMRKVVIMLQEV 967


RESULT 13
KRAF_CAEEL
ID   KRAF_CAEEL     STANDARD;      PRT;   813 AA.
AC   Q07292;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   RAF HOMOLOG SERINE/THREONINE-PROTEIN KINASE (EC 2.7.1.-).
GN   LIN-45 OR RAF-1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93247635; PubMed=8483497;
RA   Han M., Golden A., Han Y., Sternberg P.W.;
RT   "C. elegans lin-45 raf gene participates in let-60 ras-stimulated
RT   vulval differentiation.";
RL   Nature 363:133-140(1993).
CC   -!- FUNCTION: PROTEIN KINASE THAT PARTICIPATES IN THE INDUCTION OF
CC       C.ELEGANS VULVA. ACTS DOWNSTREAM OF THE RAS PROTEIN LET-60.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       MIL/RAF SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
CC       BINDING DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   -----------------------------------------------------------------------
DR   EMBL; L15347; AAA28142.1; -.
DR   PIR; S33261; S33261.
DR   HSSP; P11362; 1AGW.
DR   InterPro; IPR000719; -.
DR   InterPro; IPR002219; -.
DR   InterPro; IPR002290; -.
DR   Pfam; PF00130; DAG_PE-bind; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding; Zinc;
KW   Phorbol-ester binding.
FT   DOMAIN      171    217       PHORBOL-ESTER AND DAG BINDING.
FT   DOMAIN      481    748       PROTEIN KINASE.
FT   NP_BIND     487    495       ATP (BY SIMILARITY).
FT   BINDING     507    507       ATP (BY SIMILARITY).
FT   ACT_SITE    602    602       BY SIMILARITY.
SQ   SEQUENCE   813 AA;  90491 MW;  6376E968C6859E49 CRC64;

  Query Match             10.2%;  Score 287.5;  DB 1;  Length 813;
  Best Local Similarity   27.7%;  Pred. No. 1.4e-12;
  Matches   91;  Conservative   65;  Mismatches  127;  Indels   45;  Gaps   14;

Qy      11 PTIPYHKLADLRYLS-------RGASGTVSSARHADWRVQVAVKHLHIHTPLLDSERKDV 63
           |  |:|:  |   |        :  ||:  :    ::   ||:| |::  |   |:
Db     466 PQKPHHE--DWEILPNEFIIQYKVGSGSFGTVYRGEFFGTVAIKKLNVVDP-TPSQMAAF 522

Qy      64 LREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLR-- 121
             |  :| | |   :   :|   ||| : |:|::    ||   :|   : | | : :
Db     523 KNEVAVLKKTRHLNVLLFMGWVREPE-IAIITQWCEGSSLYRHIH--VQEPRVEFEMGAI 579

Qy     122 FRILHEIALGVNYLHNMTPPLLHHDLKTQNI-LLDNEFHVKIADFGL----SKWRMMSLS 176
             || :::||:||||:    ::| |||| || |:|:   ||| ||||    :|| :
Db     580 IDILKQVSLGMNYLHSKN--IIHRDLKTNNIFLMDDMSTVKIGDFGLATVKTKWTVNGGQ 637

Qy     177 QSRSSKSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPL 236
           | :        |:|::| ||       :  : :|:|| : :  :|:||   |: :: |
Db     638 QQQQPT-----GSILWMAPEVIRMQDDNPYTPQSDVYSFGICMYEILSSHLPYSNINNRD 692

Qy     237 QIMYSVSQGH-RPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLR 295
           ||:: | :|: ||  :   : :| |    |: | ::    : :||| |
Db     693 QILFMVGRGYLRP--DRSKIRHDTP--KSMLKLYDNCIMFDRNERPVF------------ 736

Qy     296 TFEEITFLEAVIQLKKTKLQSVSSAIHL 323
             | :  |  :|  | |: ||  : :||
Db     737 -GEVLERLRDIILPKLTRSQSAPNVLHL 763


RESULT  14
ANPB_HUMAN
ID   ANPB_HUMAN     STANDARD;      PRT;  1047 AA.
AC   P20594; Q9UQ50; O60871;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   ATRIAL NATRIURETIC PEPTIDE RECEPTOR B PRECURSOR (ANP-B) (ANPRB) (GC-B)
DE   (GUANYLATE CYCLASE) (EC 4.6.1.2) (NPR-B) (ATRIAL NATRIURETIC PEPTIDE
DE   B-TYPE RECEPTOR).
GN   NPR2 OR ANPRB.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (LONG ISOFORM).
RC   TISSUE=Brain;
RX   MEDLINE=89365195; PubMed=2570358;
RA   Chang M.S., Lowe D.G., Lewis M., Hellmiss R., Chen E., Goeddel D.V.;
```

```
RT   "Differential activation by atrial and brain natriuretic peptides of
RT   two different receptor guanylate cyclases.";
RL   Nature 341:68-72(1989).
RN   [2]
RP   SEQUENCE FROM N.A. (LONG ISOFORM).
RC   TISSUE=Blood;
RX   PubMed=10082481;
RA   Rehemudula D., Nakayama T., Soma M., Takahashi Y., Uwabo J., Sato M.,
RA   Izumi Y., Kanmatsuse K., Ozawa Y.;
RT   "Structure of the type B human natriuretic peptide receptor gene and
RT   association of a novel microsatellite polymorphism with essential
RT   hypertension.";
RL   Circ. Res. 84:605-610(1999).
RN   [3]
RP   SEQUENCE FROM N.A. (SHORT ISOFORM).
RC   TISSUE=Kidney;
RA   Hirsch J.R., Meyer M., Magert H.J., Forssmann W.G., Mollerup S.,
RA   Herter P., Weber G., Cermak R., Ankorina-Stark I., Schlatter E.,
RA   Kruhoffer M.;
RT   "cGMP-dependent and independent inhibition of a K+ conductance by
RT   natriuretic peptides. Molecular and functional studies in human
RT   proximal tubule cells.";
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. HAS GUANYLATE
CC       CYCLASE ACTIVITY ON BINDING OF ANF. SEEMS TO BE STIMULATED MORE
CC       EFFECTIVELY BY BRAIN NATRIURETIC PEPTIDE (BNP) THAN BY ANP.
CC   -!- CATALYTIC ACTIVITY: GTP = 3,'5'-CYCLIC GMP + PYROPHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC       SHORT FORM/NPR-BI; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO
CC       WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C)
CC       WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE
CC       CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.
CC   -!- SIMILARITY: SOME SIMILARITY WITH CONSERVED REGION OF CATALYTIC
CC       DOMAIN OF PROTEIN KINASES.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   -!- SIMILARITY: TO ANP-A ON THE COMPLETE SEQUENCE, AND TO ANP-C IN TH
CC       EXTRACELLULAR AND TRANSMEMBRANE DOMAINS.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AB005647; BAA81737.1; -.
DR   EMBL; AB005626; BAA81737.1; JOINED.
DR   EMBL; AB005627; BAA81737.1; JOINED.
DR   EMBL; AB005628; BAA81737.1; JOINED.
DR   EMBL; AB005629; BAA81737.1; JOINED.
DR   EMBL; AB005630; BAA81737.1; JOINED.
DR   EMBL; AB005631; BAA81737.1; JOINED.
DR   EMBL; AB005632; BAA81737.1; JOINED.
DR   EMBL; AB005633; BAA81737.1; JOINED.
DR   EMBL; AB005634; BAA81737.1; JOINED.
DR   EMBL; AB005635; BAA81737.1; JOINED.
DR   EMBL; AB005636; BAA81737.1; JOINED.
DR   EMBL; AB005637; BAA81737.1; JOINED.
DR   EMBL; AB005638; BAA81737.1; JOINED.
DR   EMBL; AB005639; BAA81737.1; JOINED.
DR   EMBL; AB005640; BAA81737.1; JOINED.
DR   EMBL; AB005641; BAA81737.1; JOINED.
DR   EMBL; AB005642; BAA81737.1; JOINED.
DR   EMBL; AB005643; BAA81737.1; JOINED.
DR   EMBL; AB005644; BAA81737.1; JOINED.
DR   EMBL; AB005645; BAA81737.1; JOINED.
DR   EMBL; AB005646; BAA81737.1; JOINED.
DR   EMBL; AJ005282; CAA06466.1; -.
DR   PIR; S05514; S05514.
```

```
DR   HSSP; Q02846; 1AWL.
DR   MIM; 108961; -.
DR   InterPro; IPR000719; -.
DR   InterPro; IPR001054; -.
DR   InterPro; IPR001170; -.
DR   InterPro; IPR001828; -.
DR   Pfam; PF01094; ANF_receptor; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR00255; NATPEPTIDER.
DR   PROSITE; PS00458; ANF_RECEPTORS; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;
KW   cGMP synthesis; Signal; Alternative splicing.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23   1047       ATRIAL NATRIURETIC PEPTIDE RECEPTOR B.
FT   DOMAIN       23    458       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    459    478       POTENTIAL.
FT   DOMAIN      479   1047       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      513    786       PROTEIN KINASE LIKE.
FT   DISULFID     75    101       BY SIMILARITY.
FT   DISULFID    439    439       INTERCHAIN (PROBABLE).
FT   DISULFID    448    448       INTERCHAIN (PROBABLE).
FT   CARBOHYD     24     24       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     35     35       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    161    161       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    195    195       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    244    244       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    277    277       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    349    349       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    964   1047       PVCAGVVGLKMPRYCLFGDTVNTASRMESNGQALKIHVSST
FT                                TKDALDELGCFQLELRGDVEMKGKGKMRTYWLLGERKGPPG
FT                                LL -> KADSHSSPSLHLSQTLPTCFFSKGQSVLGLLA
FT                                (IN SHORT ISOFORM).
FT   CONFLICT    755    755       T -> S (IN REF. 2).
SQ   SEQUENCE   1047 AA;  117021 MW;  817FB74D6B31F7EF CRC64;

  Query Match             10.2%;  Score 287.5;  DB 1;  Length 1047;
  Best Local Similarity   27.1%;  Pred. No. 2e-12;
  Matches  101;  Conservative   80;  Mismatches  115;  Indels   77;  Gaps   21;

Qy     15 YHKLADLRY-LS-RGAS-GTVSSARHADWRV----------QVAVKHLHIHTPLLDSERK 61
          ||| |  |  || ||:| |:: :| |  :::           ||:|  |::   ::  |:
Db    507 YHKGAGSRLTLSLRGSSYGSLMTA-HGKYQIFANTGHFKGNVVAIK--HVNKKRIELTRQ 563

Qy     62 DVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLR 121
           || | : :   :|:::   :| | :|  : ||||| | ||| ::|  : :  :: |  |
Db    564 -VLFELKHMRDVQFNHLTRFIGACIDPPNICIVTEYCPRGSLQDIL--ENDSINLDWMFR 620

Qy    122 FRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSS 181
          : :::::  |: :|||      |  ||: | ::|: | :|| |:||: :|         |
Db    621 YSLINDLVKGMAFLHNSIIS-SHGSLKSSNCVVDSRFVLKITDYGLASFR---------S 670

Qy    182 KSAPEGGTIIY-----MPPE----NYEPGQKSRASIKHDIYSYAVITWEVLSRKQPF--E 230
           : |:    :|       ||    |  |   :    | |:||: :|  |:  |  ||  |
Db    671 TAEPDDSHALYAKKLWTAPELLSGNPLP---TTGMQKADVYSFGIILQEIALRSGPFYLE 727

Qy    231 DV-TNPLQIMYSVSQGHRPV---------INEESLPYDIPHRARMISLIESGWAQNPDER 280
           :  :| :|:  |  | ||          :|||           :: |:|  |||:| ||
Db    728 GLDLSPKEIVQKVRNGQRPYFRPSIDRTQLNEE-----------LVLLMERCWAQDPAER 776

Qy    281 PSFLKCLIELEPVLRTFEE---ITFLE-AVIQLKK-----TKLQSVSSAIHLCDKKKMEL 331
          | |     :::  :| | :    : |:  ::::::      ||    :  :| :|:| |
Db    777 PDF----GQIKGFIRRFNKEGGTSILDNLLLRMEQYANNLEKLVEERTQAYLEEKRKAEA 832

Qy    332 SLNIPVNHGPQEE 344
           |   : |   |:
Db    833 LLYQILPHSVAEQ 845


RESULT 15
```

```
M3K9_HUMAN
ID   M3K9_HUMAN     STANDARD;      PRT;   394 AA.
AC   P80192;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   MITOGEN-ACTIVATED PROTEIN KINASE KINASE KINASE 9 (EC 2.7.1.-) (MIXED
DE   LINEAGE KINASE 1) (FRAGMENT).
GN   MAP3K9 OR MLK1 OR PRKE1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Colon epithelium;
RX   MEDLINE=93238756; PubMed=8477742;
RA   Dorow D.S., Devereux L., Dietzsch E., de Kretser T.;
RT   "Identification of a new family of human epithelial protein kinases
RT   containing two leucine/isoleucine-zipper domains.";
RL   Eur. J. Biochem. 213:701-710(1993).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN EPITHELIAL TUMOR CELL LINES OF
CC       COLONIC, BREAST AND OESOPHAGEAL ORIGIN.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       MAP KINASE KINASE KINASE SUBFAMILY.
DR   PIR; S32467; S32467.
DR   PIR; JU0229; JU0229.
DR   HSSP; P11362; 1FGI.
DR   MIM; 600136; -.
DR   InterPro; IPR000719; -.
DR   InterPro; IPR002290; -.
DR   Pfam; PF00069; pkinase; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
KW   ATP-binding.
FT   NON_TER       1      1
FT   DOMAIN        3    271       PROTEIN KINASE.
FT   NP_BIND       9     17       ATP (BY SIMILARITY).
FT   BINDING      30     30       ATP (BY SIMILARITY).
FT   ACT_SITE    127    127       BY SIMILARITY.
FT   DOMAIN      289    310       LEUCINE-ZIPPER (BY SIMILARITY).
FT   DOMAIN      324    345       LEUCINE-ZIPPER (BY SIMILARITY).
FT   DOMAIN      354    368       ARG/LYS-RICH (BASIC).
SQ   SEQUENCE   394 AA;  44975 MW;  DBE40B7D31047FD8 CRC64;

  Query Match             10.0%;  Score 284;  DB 1;  Length 394;
  Best Local Similarity   26.4%;  Pred. No. 9.1e-13;
  Matches  111;  Conservative   62;  Mismatches  125;  Indels  122;  Gaps   21;

Qy     43 QVAVKHLHIHTPLLDSER--KDVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPN 100
          :||||    | |  |  :  ::| :||::    :   |  : |:| :   | :| |:
Db     26 EVAVKAAR-HDPDEDISQTIENVRQEAKLFAMLKHPNIIALRGVCLKEPNLCLVMEFARG 84

Qy    101 GSLNELLHRKTEYPD--VAWPLRFRILHEIALGVNYLHN-MTPPLLHHDLKTQNIL---- 153
          | || :|  |   ||  | | :      :|| |:||||:    |::| |||: |||
Db     85 GPLNRVLSGKRIPPDILVNWAV------QIARGMNYLHDEAIVPIIHRDLKSSNILILQK 138

Qy    154 -----LDNEFHVKIADFGLSK-WRMMSLSQSRSSKSAPEGGTIIYMPPENYEPGQKSRAS 207
               | |:  :|| ||||:: |        |::| :   ||  :| ||       |: |
Db    139 VENGDLSNKI-LKITDFGLAREWH-------RTTKMS-AAGTYAWMAPEVIRASMFSKGS 189

Qy    208 IKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIPHRARMIS 267
             |::|| |: ||:|: : ||  : : |:: | |:      :|: :||
Db    190 ---DVWSYGVLLWELLTGEVPFRGI-DGLRVAYGVA------MNKLALPIPSTCPEPFAK 239

Qy    268 LIESGWAQNPDERPSFLKCLIEL------------------------------------- 290
          |:|  |  :|  ||||   | :|
Db    240 LMEDCWNPDPHSRPSFTNILDQLTTIEESGFFEMPKDSFHCLQDNWKHEIQEMFDQLRAK 299

Qy    291 EPVLRTFEEITFLEAVIQLKKTKLQSVSSAIHLCDKKKMELS----------LNIPVNHG 340
```

```
           |  |||:||        :| :  ||  :    |  ::: ||:          ||| ::
Db     300 EKELRTWEE--------ELTRAALQQKNQE-ELLRRREQELAEREIDILERELNIIIHQL 350

Qy     341 PQEESCGSSQLHENSGSPETSR----SLPAPQDNDFLSRKAQDCYFMKLHHCPG--NHSW 394
            ||:     :: :  |    ||     || |  :                 |||    ||
Db     351 CQEK----PRVKKRKGKFRKSRLAQPVLPFPHGHS---------------RCPGGTGSSW 391


Search completed: June 14, 2001, 08:04:51
Job time: 23 sec
```

```
                    GenCore version 4.5
             Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          June 14, 2001, 08:04:28 ; Search time 53.04 Seconds
                                           (without alignments)
                                           699.668 Million cell updates/sec

Title:           US-09-445-223-1
Perfect score:   2829
Sequence:        1 MNGEAICSALPTIPYHKLAD..........PEILVVSRSPSLNLLQNKSM 540

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters:      198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_67:*
               1:  pir1:*
               2:  pir2:*
               3:  pir3:*
               4:  pir4:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                    %
Result           Query
  No.   Score    Match Length DB  ID                   Description
-------------------------------------------------------------------------------
    1     394    13.9    656  2  I49299               receptor interacti
    2   385.5    13.6    829  2  T07406               probable protein k
    3     376    13.3    821  2  T48400               serine/threonine-p
    4     369    13.0    671  2  T09479               serine/threonine p
    5   359.5    12.7    963  2  T09911               probable serine/th
    6     359    12.7   1015  2  T00726               probable serine/th
    7   348.5    12.3    370  2  T46150               protein kinase ATN
    8     340    12.0    982  2  T06576               probable protein k
    9   335.5    11.9    736  2  T05137               protein kinase hom
   10   327.5    11.6    406  2  T52626               probable mitogen-a
   11   317.5    11.2    462  2  S29851               protein kinase 6 (
   12   316.5    11.2    776  2  T02584               hypothetical prote
   13   313.5    11.1   1584  2  T18276               protein-tyrosine k
   14     312    11.0    579  2  JC5955               transforming growt
   15     307    10.9    553  2  T04683               hypothetical prote
   16   306.5    10.8    475  2  T12955               probable protein k
   17   306.5    10.8    567  2  JC5957               transforming growt
   18   306.5    10.8    888  2  A55318               serine/threonine p
   19   306.5    10.8    888  2  JC5399               dual leucine zippe
   20     305    10.8    784  2  T45697               hypothetical prote
   21   303.5    10.7    410  2  B35670               protein-tyrosine k
   22   302.5    10.7    390  2  T01451               protein kinase hom
   23   301.5    10.7    886  2  T48544               MAP3K delta-1 prot
   24   300.5    10.6    606  2  JC5956               transforming growt
   25     300    10.6    443  2  T01182               hypothetical prote
   26   297.5    10.5    545  2  T05675               hypothetical prote
   27     295    10.4   1050  2  S45636               natriuretic-peptid
   28   292.5    10.3    668  2  JC2363               protein kinase (EC
   29   292.5    10.3   1047  1  OYRTBR               atrial natriuretic
   30     292    10.3   1338  2  T18287               protein-tyrosine k
   31   290.5    10.3    816  2  T45684               hypothetical prote
   32     290    10.3    954  1  S68178               mixed-lineage prot
   33   289.5    10.2    999  1  S27756               receptor-like prot
   34   289.5    10.2   1192  2  T48499               receptor-like prot
   35     289    10.2    855  2  T10665               hypothetical prote
   36   288.5    10.2    871  2  T45692               receptor-like prot
   37     288    10.2    884  2  T02731               serine/threonine-s
   38   287.5    10.2    813  1  S33261               protein kinase lin
   39   287.5    10.2   1047  1  OYHUBR               natriuretic peptid
   40     287    10.1    656  2  T10664               serine/threonine-s
   41     286    10.1    819  2  T45690               receptor-like prot
   42   285.5    10.1    694  2  T01134               hypothetical prote
   43     284    10.0    394  2  JU0229               mixed-lineage prot
   44     283    10.0    356  2  T50811               ser/thr specific p
   45     283    10.0    402  2  T51791               ser/thr specific p

                              ALIGNMENTS

RESULT   1
I49299
receptor interacting protein RIP - mouse
C;Species: Mus musculus (house mouse)
C;Date: 09-Mar-1996 #sequence_revision 09-Mar-1996 #text_change 05-Nov-1999
C;Accession: I49299
R;Stanger, B.Z.; Leder, P.; Lee, T.H.; Kim, E.; Seed, B.
Cell 81, 513-523, 1995
A;Title: RIP: a novel protein containing a death domain that interacts with Fas/APO-1
A;Reference number: A56913; MUID:95277838
A;Accession: I49299
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-656 <RES>
A;Cross-references: EMBL:U25995; NID:g829618; PIDN:AAB60487.1; PID:g829619
C;Genetics:
A;Gene: RIP
C;Superfamily: protein kinase homology
F;15-293/Domain: protein kinase homology <KIN>

  Query Match             13.9%;  Score 394;  DB 2;  Length 656;
  Best Local Similarity   31.0%;  Pred. No. 1.1e-15;
  Matches  131;  Conservative   66;  Mismatches  144;  Indels   82;  Gaps   19;

Qy      24 LSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSERKDV-LREAEILHKARFSYIFPIL 82
           |  |  | ||   |      :  |   ::|    :|  :| | | :::|: | | :  :|
Db      23 LDSGGFGKVSLCYHRSHGFVILKK---VYTGPNRAEYNEVLLEEGKMMHRLRHSRVVKLL 79

Qy      83 GICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPL 142
           ||  |     :| |||  |:|  :|  ||:  ||   |: ||: |   |: |||:    :
Db      80 GIIIEEGNYSLVMEYMEKGNLMHVL--KTQI-DVPLSLKGRIIVEAIEGMCYLHD--KGV 134

Qy     143 LHHDLKTQNILLDNEFHVKIADFGLSKWRMMS-LSQSRSSKSAP--------EGGTIIYM 193
           :| ||| :|||:| :||:|||| |:: ::  | |:: : :|            |||: ||
Db     135 IHKDLKPENILVDRDFHIKIADLGVASFKTWSKLTKEKDNKQKEVSSTTKKNNGGTLYYM 194

Qy     194 PPE-----NYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVSQGHRP 248
            ||     | :| :||      |:||: :: | : ::|:|:|:|    | :  :  |:||
Db     195 APEHLNDINAKPTEKS------DVYSFGIVLWAIFAKKEPYENVICTEQFVICIKSGNRP 248

Qy     249 VINE--ESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEP-VLRTFEEITFLEA 305
            : |  |  | :|      |||:|  |   |::||:||    |  |  |  |||    |
Db     249 NVEEILEYCPREI------ISLMERCWQAIPEDRPTFLGIEEEFRPFYLSHFEEYV-EED 301

Qy     306 VIQLKK------TKLQSVSSAIHLC---------DKKKMELSLNIPVNHGPQEESCGSSQ 350
           |  |||        || : |  | |          ::   |  :  :  || |||  ||
Db     302 VASLKKEYPDQSPVLQRMFSLQHDCVPLPPSRSNSEQPGSLHSSQGLQMGPVEESWFSS- 360

Qy     351 LHENSGSPETSRSLPAPQDNDFLSRKAQDCYFMKLHHCPGNHSWDSTISGSQRAAFCDHK 410
                 |||       ||| :  | :|      ||      |::           | : :
Db     361 ------SPE------YPQDENDRSVQA------KLQEEASYHAF---------GIFAEKQ 393
```

```
Qy    411 TTP 413
          | |
Db    394 TKP 396

RESULT   2
T07406
probable protein kinase - tomato
C;Species: Lycopersicon esculentum (tomato)
C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 08-Oct-1999
C;Accession: T07406
R;Wang, Y.
submitted to the EMBL Data Library, May 1997
A;Reference number: Z16016
A;Accession: T07406
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-829 <WAN>
A;Cross-references: EMBL:Y13273; NID:e1050452; PIDN:CAA73722.1; PID:e334294
A;Experimental source: strain UC82B; sub_species Mill
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo


  Query Match             13.6%;  Score 385.5;  DB 2;  Length 829;
  Best Local Similarity   33.2%;  Pred. No. 4.7e-15;
  Matches  104;  Conservative   56;  Mismatches  108;  Indels   45;  Gaps   12;

Qy      3 GEAICSALPT-----------IPYHKLADLRYLSRGASGTVSSARHADWR-VQVAVKHL- 49
          |  : ||:|:           ||:: |  :  :  |: |||      ||    |||| |
Db    529 GGHVVSAIPSEELDLDVEEFNIPWNDLILMEKIGAGSFGTV---HRGDWHGSDVAVKILM 585

Qy     50 --HIHTPLLDSERKDVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELL 107
              |   |    |: |||  |: : |   |   :|   :|  | |||||:  |||  ||
Db    586 EQDFHAERL----KEFLREVAIMKRLRHPNIVLFMGAVIQPPNLSIVTEYLSRGSLYRLL 641

Qy    108 HRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGL 167
          |:      :    |  : :::| |:||||   ||::| |||: |:|:| :: ||| ||||
Db    642 HKPGAREVLDERRRLCMAYDVANGMNYLHKRNPPIVHRDLKSPNLLVDKKYTVKICDFGL 701

Qy    168 SKWRMMSLSQSRSSKSAPEGGTIIYMPPE--NYEPGQKSRASIKHDIYSYAVITWEVLSR 225
          |:::  :   |:::   ||     :| ||    ||  :     | |:||: || ||: :
Db    702 SRFKANTFLSSKTAAGTPE-----WMAPEVIRDEPSNE-----KSDVYSFGVILWELATL 751

Qy    226 KQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIPH--RARMISLIESGWAQNPDERPSF 283
          :||:  : || |:: :|        | : |  |||     ::  :||: ||  | :||||
Db    752 QQPWNKL-NPPQVIAAVG------FNRKRL--DIPSDLNPQVAIIIEACWANEPWKRPSF 802

Qy    284 LKCLIELEPVLRT 296
             :  | | |::
Db    803 STIMDMLRPHLKS 815


RESULT   3
T48400
serine/threonine-protein kinase ctrl - Arabidopsis thaliana
N;Alternate names: protein F17C15.150
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 17-Nov-2000
C;Accession: T48400; A45178
R;Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.;
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24492
A;Accession: T48400
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-821 <BEV>
A;Cross-references: EMBL:AL162506
A;Experimental source: cultivar Columbia; BAC clone F17C15
R;Kieber, J.J.; Rothenberg, M.; Roman, G.; Feldmann, K.A.; Ecker, J.R.
Cell 72, 427-441, 1993
A;Title: CTR1, a negative regulator of the ethylene response pathway in Arabidopsis, enc
A;Reference number: A45178; MUID:93161417
```

```
A;Contents: Columbia
A;Accession: A45178
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-468,470-821 <KIE>
A;Note: sequence extracted from NCBI backbone (NCBIP:124878)
C;Genetics:
A;Map position: 5
A;Introns: 216/3; 317/2; 364/3; 399/2; 468/1; 560/1; 573/3; 596/3; 630/2; 658/3; 691/
A;Note: F17C15.150
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
C;Keywords: ATP
F;549-812/Domain: protein kinase homology <KIN>
F;557-565/Region: protein kinase ATP-binding motif


  Query Match             13.3%;  Score 376;  DB 2;  Length 821;
  Best Local Similarity   30.7%;  Pred. No. 1.6e-14;
  Matches   98;  Conservative   60;  Mismatches  111;  Indels   50;  Gaps   12;

Qy      1 MNGEAICSALPT--------------IPYHKLADLRYLSRGASGTVSSARHADWR-VQVA 45
          ||   |   :|               ||:    ||    :  :|:  :   |:|    ||
Db    520 MNAPPISQPVPNRANRELGLDGDDMDIPW---CDLNIKEKIGAGSFGTVHRAEWHGSDVA 576

Qy     46 VKHL---HIHTPLLDSER-KDVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNG 101
          || |     |     :||  : |||  |: : |   |   :|   :|  | |||||:  |
Db    577 VKILMEQDFH-----AERVNEFLREVAIMKRLRHPNIVLFMGAVTQPPNLSIVTEYLSRG 631

Qy    102 SLNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVK 161
          ||  |||:      :    |  : :::| |:|||||  ||::| |||: |:|:| :: ||
Db    632 SLYRLLHKSGAREQLDERRRLSMAYDVAKGMNYLHNRNPPIVHRDLKSPNLLVDKKYTVK 691

Qy    162 IADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPE--NYEPGQKSRASIKHDIYSYAVIT 219
          : |||||: :  :   |:|:   ||     :| ||    ||  :     | |:||: ||
Db    692 VCDFGLSRLKASTFLSSKSAAGTPE-----WMAPEVLRDEPSNE-----KSDVYSFGVIL 741

Qy    220 WEVLSRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIPH--RARMISLIESGWAQNP 277
          ||: : :||: :: || |:: :|    : :        :||     :: ::||  |   |
Db    742 WELATLQQPWGNL-NPAQVVAAVGFKCKRL--------EIPRNLNPQVAAIIEGCWTNEP 792

Qy    278 DERPSFLKCLIELEPVLRT 296
           :||||   :  | |::::
Db    793 WKRPSFATIMDLLRPLIKS 811


RESULT   4
T09479
serine/threonine protein kinase (EC 2.7.1.-) RIP - human
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 17-Nov-2000
C;Accession: T09479; I38992
R;Huang, J.; Hsu, H.; Baichwal, V.R.; Goeddel, D.V.
submitted to the EMBL Data Library, August 1998
A;Reference number: Z16685
A;Accession: T09479
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-671 <HUA>
A;Cross-references: EMBL:U50062; NID:g3426026; PID:g3426027
R;Stanger, B.Z.; Leder, P.; Lee, T.H.; Kim, E.; Seed, B.
Cell 81, 513-523, 1995
A;Title: RIP: a novel protein containing a death domain that interacts with Fas/APO-1
A;Reference number: A56913; MUID:95277838
A;Accession: I38992
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 300-513,'S',515-671 <RES>
A;Cross-references: EMBL:U25994; NID:g829616; PIDN:AAC50137.1; PID:g829617
C;Genetics:
A;Gene: RIP
C;Keywords: ATP binding; phosphotransferase
```

```
  Query Match             13.0%;  Score 369;  DB 2;  Length 671;
  Best Local Similarity   28.8%;  Pred. No. 3.3e-14;
  Matches  112;  Conservative   67;  Mismatches  146;  Indels   64;  Gaps   12;

Qy      24 LSRGASGTVSSARHADWRVQ-VAVKHLHIHTPLLDSERKDVLREAEILHKARFSYIFPIL 82
           |  |  | ||   |   | | : :       |      : :| ||::::: | | :  :|
Db      23 LDSGGFGKVSLCFH---RTQGLMIMKTVYKGPNCIEHNEALLEEAKMMNRLRHSRVVKLL 79

Qy      83 GICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPL 142
           |:  |     :| |||  |:|  :|  :   |      ::||: ||  |: |||     :
Db      80 GVIIEEGKYSLVMEYMEKGNLMHVLKAEMSTP---LSVKGRIILEIIEGMCYLHG--KGV 134

Qy     143 LHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPE--------GGTIIYMP 194
           :| ||| :|||:||:||:|||| ||: ::| |   :       |        |||: ||
Db     135 IHKDLKPENILVDNDFHIKIADLGLASFKMWSKLNNEEHNELREVDGTAKKNGGTLYYMA 194

Qy     195 PE-----NYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVSQGHRPV 249
           ||     | :| :||      |:||:||: | : : |:|:|:     |::  :  |:||
Db     195 PEHLNDVNAKPTEKS------DVYSFAVVLWAIFANKEPYENAICEQQLIMCIKSGNRPD 248

Qy     250 INE--ESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEITFLEAVI 307
           :::  |  | :|      |||::  |  ||: ||:|     :  |   :  | :  | |
Db     249 VDDITEYCPREI------ISLMKLCWEANPEARPTFPGIEEKFRPFYLSQLEESVEEDVK 302

Qy     308 QLKK---------TKLQSV--------SSAIHLCDKKKMELSLNIPVNHGPQEESCGSSQ 350
            |||          ::||:        ||  :   ::   |  :  :  || |||  :
Db     303 SLKKEYSNENAVVKRMQSLQLDCVAVPSSRSNSATEQPGSLHSSQGLGMGPVEESWFAPS 362

Qy     351 LHENSGSPETSRSLPAPQDNDFLSRKAQD 379
           |             |  ::    |  | ||
Db     363 LEH-----------PQEENEPSLQSKLQD 380

RESULT   5
T09911
probable serine/threonine-specific protein kinase (EC 2.7.1.-) T22A6.310 - Arabidopsis t
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 15-Oct-1999
C;Accession: T09911
R;Bevan, M.; Zimmermann, W.; Grueneisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; May
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z16896
A;Accession: T09911
A;Molecule type: DNA
A;Residues: 1-963 <BEV>
A;Cross-references: EMBL:AL078637; GSPDB:GN00062; ATSP:T22A6.310
A;Experimental source: cultivar Columbia; BAC clone T22A6
C;Genetics:
A;Gene: ATSP:T22A6.310
A;Map position: 4
A;Introns: 286/3; 386/2; 434/3; 473/2; 543/1; 678/1; 691/3; 719/3; 753/2; 788/3; 821/3;
C;Keywords: phosphotransferase; protein kinase

  Query Match             12.7%;  Score 359.5;  DB 2;  Length 963;
  Best Local Similarity   31.7%;  Pred. No. 1.8e-13;
  Matches   93;  Conservative   60;  Mismatches  105;  Indels   35;  Gaps   11;

Qy      20 DLRYLSRGASGTVSSARHADWR-VQVAVKHLHIHTPLLDSERKDVLRE-----AEILHKA 73
           :|    |  :|:  :   |:|    |||| | |     | : :: |||       |: :
Db     668 ELHIKERVGAGSFGTVHRAEWHGSDVAVKILSIQ-DFHDDQFREFLREVCKQAVAIMKRV 726

Qy      74 RFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKT--EYPDVAWPLRFRI-----LH 126
           |   :   :|   |   | |:|||:| |||  |:||    |  |    ||  :     :
Db     727 RHPNVVLFMGAVTERPRLSIITEYLPRGSLFRLIHRPASGELLDQRRRLRMALDVVCAIP 786

Qy     127 EIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPE 186
             | |:|||| : ||::| |||: |:|:|  : ||: |||||:::  :   |:|    ||
Db     787 HYAKGLNYLHCLNPPVVHWDLKSPNLLVDKNWTVKVCDFGLSRFKANTFIPSKSVAGTPE 846

Qy     187 GGTIIYMPPE--NYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVS- 243
                :| ||    ||  :     | |:||: |: ||::: :||:  : :| |:: :|:
Db     847 -----WMAPEFLRGEPTNE-----KSDVYSFGVVLWELITLQQPWNGL-SPAQVVGAVAF 895

Qy     244 QGHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRT 296
           |  | :|   : |        ::||:|: ||  | :||:|   :  |: :|::
Db     896 QNRRLIIPPNTSPV-------LVSLMEACWADEPSQRPAFGSIVDTLKKLLKS 941

RESULT   6
T00726
probable serine/threonine-specific protein kinase (EC 2.7.1.-) F22O13.21 - Arabidopsi
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 22-Oct-1999
C;Accession: T00726
R;Shinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.;
eologis, A.; Ecker, J.R.
submitted to the EMBL Data Library, April 1998
A;Description: Genomic sequence for Arabidopsis thaliana BAC F22O13.
A;Reference number: Z14200
A;Accession: T00726
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1015 <SHI>
A;Cross-references: EMBL:AC003981; NID:g3063438; PID:g3063459; GSPDB:GN00059; ATSP:F2
A;Experimental source: cultivar Columbia
C;Genetics:
A;Gene: ATSP:F22O13.21
A;Map position: 1
A;Introns: 71/3; 219/3; 360/3; 395/2; 760/1; 773/3; 796/3; 830/2; 856/3; 889/3; 912/3
C;Keywords: phosphotransferase; protein kinase

  Query Match             12.7%;  Score 359;  DB 2;  Length 1015;
  Best Local Similarity   32.9%;  Pred. No. 2.1e-13;
  Matches   96;  Conservative   53;  Mismatches  105;  Indels   38;  Gaps   11;

Qy      13 IPYHKLADLRYLSRGASGTVSSARHADWR-VQVAVKHLHIHTPLLDSE-----RKDVLRE 66
           ||:: |     :  |: | |    ||||   :||||        || :       :   |
Db     746 IPWNDLVIAERIGLGSYGEV---YHADWHGTEVAVK------KFLDQDFSGAALAEFRSE 796

Qy      67 AEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILH 126
             |: : |   :   ||    |  | ||||::| |||  :|||   : |     | ::
Db     797 VRIMRRLRHPNVVFFLGAVTRPPNLSIVTEFLPRGSLYRILHRPKSHIDER--RRIKMAL 854

Qy     127 EIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPE 186
           ::|:|:| ||  || ::| |||| |:|:|| ::||: |||||: :  :   |:|:   ||
Db     855 DVAMGMNCLHTSTPTIVHRDLKTPNLLVDNNWNVKVGDFGLSRLKHNTFLSSKSTAGTPE 914

Qy     187 GGTIIYMPPE--NYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVS- 243
                :| ||    ||  :     | |:||: || ||: : : |:  : ||:|:: :|
Db     915 -----WMAPEVLRNEPSNE-----KCDVYSFGVILWELATLRLPWRGM-NPMQVVGAVGF 963

Qy     244 QGHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLR 295
           |  |  | :|  |        :  :|   |  :|: |||| :    |:|: |
Db     964 QNRRLEIPKELDPV-------VGRIILECWQTDPNLRPSFAQLTEVLKPLNR 1008

RESULT   7
T46150
protein kinase ATN1-like protein - Arabidopsis thaliana
N;Alternate names: protein T3A5.110
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 17-Mar-2000
C;Accession: T46150; T08394
R;Bloecker, H.; Mewes, H.W.; Mayer, K.F.X.; Lemcke, K.; Schueller, C.; Quetier, F.; S
submitted to the Protein Sequence Database, December 1999
A;Reference number: Z23024
A;Accession: T46150
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-370 <BLO>
A;Cross-references: EMBL:AL132979
A;Experimental source: cultivar Columbia; BAC clone T3A5
R;Quetier, F.; Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Salan
submitted to the Protein Sequence Database, May 1999
```

```
A;Reference number: Z16409
A;Accession: T08394
A;Molecule type: DNA
A;Residues: 1-370 <QUE>
A;Cross-references: EMBL:AL049862; GSPDB:GN00061; ATSP:F18B3.10
A;Experimental source: cultivar Columbia; BAC clone F18B3
C;Genetics:
A;Gene: ATSP:F18B3.10
A;Map position: 3
A;Introns: 54/2; 99/3; 164/1; 204/3; 257/3
A;Note: T3A5.110
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote

  Query Match             12.3%;  Score 348.5;  DB 2;  Length 370;
  Best Local Similarity   29.2%;  Pred. No. 2.4e-13;
  Matches   94;  Conservative   63;  Mismatches  106;  Indels   59;  Gaps   13;

Qy      44 VAVKHLHIHT--PLLDSERKDVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNG 101
           |||| :   |   :  : :|   :|  :| | :   |   :| | ||: : |||| :  |
Db      61 VAVKIMDPSTTSAVTKAHKKTFQKEVLLLSKMKHDNIVKFVGACIEPQLI-IVTELVEGG 119

Qy     102 SLNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEF-HV 160
           :|   :| :    |:   | | :  :|:  : ::|  :  ::| ||  :|:|:  :  ||
Db     120 TLQRFMHSRPGPLDLKMSLSFAL--DISRAMEFVH--SNGIIHRDLNPRNLLVTGDLKHV 175

Qy     161 KIADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPENYEP-----GQKSRASIKHDIYSY 215
           |:||||:::       ::|   :  | ||  :| || | |     |:|     | ||||:
Db     176 KLADFGIAR------EETRGGMTC-EAGTSKWMAPEVYSPEPLRVGEKKEYDHKADIYSF 228

Qy     216 AVITWEVLSRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIPHRARMISLIESGWAQ 275
           |:: |:::: ::|| || | | : | |||| ||:: :     |:      : ::|| |||
Db     229 AIVLWQLVTNEEPFPDVPNSLFVPYLVSQGRRPILTKTP---DV-----FVPIVESCWAQ 280

Qy     276 NPDERPSFLKCLIELEPVLRTFEEITFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLNI 335
           :|| || | :  : |  :||                 :: | ||           :   :
Db     281 DPDARPEFKEISVMLTNLLR-----------------RMSSDSS-----------IGTTL 312

Qy     336 P---VNHGPQEESCGSSQLHEN 354
           |      |  |||  |  | |:
Db     313 PDGEAYEGEMEESENSPLLQEH 334

RESULT   8
T06576
probable protein kinase TCTR2 - tomato
C;Species: Lycopersicon esculentum (tomato)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 08-Oct-1999
C;Accession: T06576
R;Hackett, R.M.
submitted to the EMBL Data Library, March 1998
A;Reference number: Z15770
A;Accession: T06576
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-982 <HAC>
A;Cross-references: EMBL:AJ005077; NID:e1296722; PIDN:CAA06334.1; PID:e1296723
A;Experimental source: cultivar Ailsa craig
C;Genetics:
A;Gene: TCTR2

  Query Match             12.0%;  Score 340;  DB 2;  Length 982;
  Best Local Similarity   30.8%;  Pred. No. 2.5e-12;
  Matches   90;  Conservative   56;  Mismatches  108;  Indels   38;  Gaps   11;

Qy      13 IPYHKLADLRYLSRGASGTVSSARHADWR-VQVAVKHLHIHTPLLDSE-----RKDVLRE 66
           ||:  |     :  |: | |    ||||   :||||        || :       :  ||
Db     696 IPWEDLVIGERIGLGSYGEV---YHADWNGTEVAVK------KFLDQDFSGAALAEFKRE 746

Qy      67 AEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILH 126
             |: : |   :   :|    |  | |:||::| |||  ::||     |    ::  :
Db     747 VRIMRRLRHPNVVRFMGAITRPPHLSIITEFLPRGSLYRIIHRPHFQIDERQKIKMAL-- 804

Qy     127 EIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPE 186
           ::| |:: ||   | ::| |||: |:|:| :::||: |||||: :  :   |:|:   ||
Db     805 DVAKGMDCLHTSNPTIVHRDLKSPNLLVDTDWNVKVCDFGLSRLKHNTFLSSKSTAGTPE 864

Qy     187 GGTIIYMPPE--NYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVS- 243
                :| ||    ||  :     | ||||: || ||: : : |:  : ||:|:: :|
Db     865 -----WMAPEVLRNEPSNE-----KCDIYSFGVILWELATLRLPWSGM-NPMQVVGAVGF 913

Qy     244 QGHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLR 295
           |  |  | :|  |        :  :|   |  :|: |||| :  : | |: |
Db     914 QNKRLEIPKELDPI-------VARIIWECWQTDPNLRPSFAQLTVALTPLQR 958

RESULT   9
T05137
protein kinase homolog F7H19.240 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C;Accession: T05137
R;Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.;
submitted to the Protein Sequence Database, July 1998
A;Reference number: Z15399
A;Accession: T05137
A;Molecule type: DNA
A;Residues: 1-736 <BEV>
A;Cross-references: EMBL:AL031018
A;Experimental source: cultivar Columbia; BAC clone F7H19
C;Genetics:
A;Map position: 4
A;Introns: 116/2; 260/3; 302/3; 477/1; 490/3; 513/3; 547/2; 573/3; 606/3; 629/3; 646/
A;Note: F7H19.240

  Query Match             11.9%;  Score 335.5;  DB 2;  Length 736;
  Best Local Similarity   32.9%;  Pred. No. 3.2e-12;
  Matches   79;  Conservative   50;  Mismatches   76;  Indels   35;  Gaps   10;

Qy      65 REAEILHKARFSYIFPILG-ICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPL--- 120
           :|  |: | |   :   :| :| | |   |: |||| ||| ::||   :      ||
Db     512 KEINIMKKLRHPNVLLFMGAVCTE-EKSAIIMEYMPRGSLFKILHNTNQ------PLDKK 564

Qy     121 -RFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSR 179
            | |:  ::| |:||||   ||::| |||: |:|:|  ::||: |||||||:  :   ::
Db     565 RRLRMALDVARGMNYLHRRNPPIVHRDLKSSNLLVDKNWNVKVGDFGLSKWKNATFLSTK 624

Qy     180 SSKSAPEGGTIIYMPPE--NYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQ 237
           | |  |:     :| ||    ||  :     | |::|: || ||:::   |: |  | :|
Db     625 SGKGTPQ-----WMAPEVLRSEPSNE-----KCDVFSFGVILWELMTTLVPW-DRLNSIQ 673

Qy     238 IMYSVSQGHRPVINEESLPYDIPH--RARMISLIESGWAQNPDERPSFLKCLIELEPVLR 295
           ::  |    | :        |:|     |: |:|:  |  :| :|||| : : ::  : |
Db     674 VVGVVGFMDRRL--------DLPEGLNPRIASIIQDCWQTDPAKRPSFEELISQMMSLFR 725

RESULT  10
T52626
probable mitogen-activated protein kinase MAP3K delta-1 [imported] - Arabidopsis thal
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 08-Dec-2000
C;Accession: T52626
R;Jouannic, S.; Hamal, A.; Leprince, A.S.; Tregear, J.W.; Kreis, M.; Henry, Y.
Gene 229, 171-81, 1999
A;Title: Characterisation of novel plant genes encoding MEKK/STE11 and RAF-related pr
A;Reference number: Z24447; MUID:99196996
A;Accession: T52626
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-406 <JOU>
A;Cross-references: EMBL:Y14199; NID:g2253009; PIDN:CAA74591.1; PID:g2253010
A;Experimental source: cultivar Columbia
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
```

```
  Query Match             11.6%;  Score 327.5;  DB 2;  Length 406;
  Best Local Similarity   32.1%;  Pred. No. 4.5e-12;
  Matches   90;  Conservative   50;  Mismatches  107;  Indels   33;  Gaps   11;

Qy      20 DLRYLSRGASGTVSSARHADWR-VQVAVKHLHIHTPLLDSE-RKDVL----REAEILHKA 73
           ||:   |   |:      |:|   :||||        || :   | |     | ||: :
Db     134 DLQIGERIGIGSYGEVYRAEWNGTEVAVK------KFLDQDFSGDALTQFKSEIEIMLRL 187

Qy      74 RFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILHEIALGVN 133
           |   :   :|    |    |:||::| |||  ||||     |     | |:  ::| |:|
Db     188 RHPNVVLFMGAVTRPPNFSILTEFLPRGSLYRLLHRPNHQLDE--KRRMRMALDVAKGMN 245

Qy     134 YLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPEGGTIIYM 193
           |||   | ::| |||: |:|:|  : ||: |||||: :  :   |:|:   ||     :|
Db     246 YLHTSHPTVVHRDLKSPNLLVDKNWVVKVCDFGLSRMKHHTYLSSKSTAGTPE-----WM 300

Qy     194 PPE--NYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVSQGHRPVIN 251
            ||    ||     |: | |:||: || ||: : : |:: : ||:|:: :|   :| :
Db     301 APEVLRNEP-----ANEKCDVYSFGVILWELATSRVPWKGL-NPMQVVGAVGFQNRRL-- 352

Qy     252 EESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELE 291
              :| ||     :  :|   |   |  |||| : :  |:
Db     353 --EIPDDID--LTVAQIIRECWQTEPHLRPSFTQLMQSLK 388

RESULT  11
S29851
protein kinase 6 (EC 2.7.1.-) - soybean
C;Species: Glycine max (soybean)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999
C;Accession: S29851; S27760
R;Feng, X.H.; Zhao, Y.; Bottino, P.J.; Kung, S.
Biochim. Biophys. Acta 1172, 200-204, 1993
A;Title: Cloning and characterization of a novel member of protein kinase family from sc
A;Reference number: S29851; MUID:93176812
A;Accession: S29851
A;Molecule type: mRNA
A;Residues: 1-462 <FEN>
A;Cross-references: EMBL:M67449; NID:g170046; PIDN:AAA34002.1; PID:g170047
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C;Keywords: ATP; phosphotransferase
F;154-419/Domain: protein kinase homology <KIN>

  Query Match             11.2%;  Score 317.5;  DB 2;  Length 462;
  Best Local Similarity   28.1%;  Pred. No. 2e-11;
  Matches  100;  Conservative   60;  Mismatches  121;  Indels   75;  Gaps   17;

Qy      28 ASGTVSSARHADWRVQ-VAVKHLHI-----HTPLLDSERKDVLREAEILHKARFSYIFPI 81
           | |  |   |  :: : |||| : :     :  |     |  :||  :| :     :
Db     163 AHGAHSRLYHGVYKDEAVAVKIIMVPEDDGNALASRLEKQFIREVTLLSRLHHQNVIKF 222

Qy      82 LGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPP 141
              | :|    |:|||:  |||   || | |:  ::         :|| |: |:|:
Db     223 SAACRKPPVYCIITEYLAEGSLRAYLH-KLEHQTISLQKLIAFALDIARGMEYIHSQ--G 279

Qy     142 LLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPENYEPG 201
           ::| ||| :|||:: : |:||||||:      :  ::     | : ||  :| ||  :
Db     280 VIHRDLKPENILINEDNHLKIADFGI------ACEEASCDLLADDPGTYRWMAPEMIK-- 331

Qy     202 QKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYS-VSQGHRPVINEESLPYDIP 260
           :||    | |:||: :| ||:|:   |:||: ||:|  :: |::  ||:|     | : |
Db     332 RKSYGK-KVDVYSFGLILWEMLTGTIPYEDM-NPIQAAFAVVNKNSRPII-----PSNCP 384

Qy     261 HRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTFEEITFLEAVIQL--------KKT 312
              |  :|||  |:  ||:|| |     ::  :|  ||     :  :  |        ||
Db     385 PAMR--ALIEQCWSLQPDKRPEFW----QVVKILEQFESSLASDGTLSLVPNPCWDHKK- 437

Qy     313 KLQSVSSAIHLCDKKKMELSLNIPVNHGPQEESCGSSQLHENSGSPETSRSLPAPQ 368
                   :|   |             ||         ||:|||        | |:
Db     438 ------GLLHWIQKL------------GP---------LHQNSG--------PVPK 458
```

```
RESULT  12
T02584
hypothetical protein T16B24.18 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 17-Mar-2000
C;Accession: T02584
R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Syke
submitted to the EMBL Data Library, August 1998
A;Description: Arabidopsis thaliana chromosome II BAC T16B24 genomic sequence.
A;Reference number: Z14679
A;Accession: T02584
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-776 <ROU>
A;Cross-references: EMBL:AC004697; NID:g3402671; PID:g3402686
A;Experimental source: cultivar Columbia
C;Genetics:
A;Map position: 2
A;Note: T16B24.18
C;Superfamily: S-receptor kinase; protein kinase homology; S-locus-specific glycoprot

  Query Match             11.2%;  Score 316.5;  DB 2;  Length 776;
  Best Local Similarity   30.2%;  Pred. No. 4.3e-11;
  Matches   87;  Conservative   57;  Mismatches  105;  Indels   39;  Gaps    8;

Qy      16 HKLADLRYLSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSERKDVLREAEILHKARF 75
           |   :   | ||: | |  |  :|  : ||||  :  | :: |  :    | ||| | |
Db     517 HGFKEFNELGRGSFGFVYKAVLSD-GIHVAVKRANAAT-IIHSNNRGFESELEILCKIRH 574

Qy      76 SYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILHEIALGVNYL 135
           : |  :|| |:|     :| ||||:|:|:: ||      |  | :| :|: : | |::||
Db     575 NNIVNLLGYCSEMGERLLVYEYMPHGTLHDHLHGDLSQLD--WSMRLKIMLQAARGLDYL 632

Qy     136 HN-MTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMP 194
           || : ||::| |:|| ||||| |   :||||||     :| ::  || |  ||
Db     633 HNEVDPPIIHRDVKTSNILLDGEMCARIADFGL-----VSSNERDSSNSDREG------- 680

Qy     195 PENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQI----MYSVSQGHRPVI 250
                           |:| :  ::  |:|| ::  :  ::|  |    :  : :|    |
Db     681 ----------------DVYDFGIVLLEILSGRKAIDRESDPAGIAEWAVPLIRKGKAAAI 724

Qy     251 NEES--LPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRT 296
            : :  || ::    ::  | |    :| :|||:    |  |: ::::
Db     725 IDRNICLPRNVEPLLKLAELAELAVRENSNERPNIRNILCFLDLIVKS 772

RESULT  13
T18276
protein-tyrosine kinase (EC 2.7.1.112) 1 - slime mold (Dictyostelium discoideum)
C;Species: Dictyostelium discoideum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Nov-2000
C;Accession: T18276; A35670
R;Nuckolls, G.H.; Osherov, N.; Loomis, W.F.; Spudich, J.A.
Development 122, 3295-3305, 1996
A;Title: The Dictyostelium dual-specificity kinase splA is essential for spore differ
A;Reference number: Z18852; MUID:97053827
A;Accession: T18276
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1584 <NUC>
A;Cross-references: EMBL:U32174; NID:g974333; PID:g974334; PIDN:AAB41125.1
R;Tan, J.L.; Spudich, J.A.
Mol. Cell. Biol. 10, 3578-3583, 1990
A;Title: Developmentally regulated protein-tyrosine kinase genes in Dictyostelium dis
A;Reference number: A35670; MUID:90287147
A;Accession: A35670
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 'R',1249-1434,'L',1436-1584 <TAN>
A;Cross-references: GB:M33785; NID:g167775; PIDN:AAA33202.1; PID:g167776
A;Note: the authors translated the codon TAT for residue 271519 as Thr
```

```
C;Genetics:
A;Gene: splA
A;Introns: 47/3; 72/2
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C;Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase
F;1287-1566/Domain: protein kinase homology <KIN>
F;1295-1303/Region: protein kinase ATP-binding motif


  Query Match             11.1%;  Score 313.5;  DB 2;  Length 1584;
  Best Local Similarity   28.1%;  Pred. No. 1.5e-10;
  Matches   91;  Conservative   59;  Mismatches  119;  Indels   55;  Gaps   10;

Qy     13 IPYHKLADLRYLSRGASGTVSSARHADWR-VQVAVKHLHIHTPLLDSERKDVLREAEILH 71
          | :::|   : : :|  | |   :   ||   ||:| ::       |       |  ||
Db   1284 IDFNELEFGQTIGKGFFGEV---KRGYWRETDVAIKIIYRDQFKTKSSLVMFQNEVGILS 1340

Qy     72 KARFSYIFPILGICNE--PEFLGIVTEYMPNGSLNELLHRKTEYPDVAWPLRFRILHEIA 129
          | |   :   || |     :   ||||:|  ||| : |       :    :| ::  :||
Db   1341 KLRHPNVVQFLGACTAGGEDHHCIVTEWMGGGSLRQFLTDHFNLLEQNPHIRLKLALDIA 1400

Qy    130 LGVNYLHNMTPPLLHHDLKTQNILLDN-------------EFHVKIADFGLSKWRMMSLS 176
           |:||||  |||:|| || ::|||||:             :   ||:|||||:     |
Db   1401 KGMNYLHGWTPPILHRDLSSRNILLDHNIDPKNPVVSSRQDIKCKISDFGLSR-----LK 1455

Qy    177 QSRSSKSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPL 236
          : ::|:     | | || ||  ::    |    | |:||| :: :|:|:  :|  :|:  |:
Db   1456 KEQASQMTQSVGCIPYMAPEVFKGDSNSE---KSDVYSYGMVLFELLTSDEPQQDM-KPM 1511

Qy    237 QIMY-SVSQGHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLR 295
          :: : :  : :||       |  :    ::   ::   |  ||| ||
Db   1512 KMAHLAAYESYRP-------PIPLTTSSKWKEILTQCWDSNPDSRP------------- 1550

Qy    296 TFEEITFLEAVIQLKKTKLQSVSS 319
               || : :: ||: : | |||
Db   1551 -----TFKQIIVHLKEMEDQGVSS 1569


RESULT 14
JC5955
transforming growth factor-beta activated kinase (EC 2.7.-.-) 1a - human
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: JC5955
R;Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.
Biochem. Biophys. Res. Commun. 243, 545-549, 1998
A;Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-in
A;Reference number: JC5955; MUID:98153801
A;Accession: JC5955
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-579 <SAK>
A;Cross-references: DDBJ:AB009356; NID:g2924623; PIDN:BAA25025.1; PID:g2924624
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C;Keywords: phosphotransferase


  Query Match             11.0%;  Score 312;  DB 2;  Length 579;
  Best Local Similarity   23.5%;  Pred. No. 5.5e-11;
  Matches  136;  Conservative   97;  Mismatches  246;  Indels  100;  Gaps   22;

Qy     13 IPYHKLADLRYLSRGASGTVSSARHADWRVQ-VAVKHLHIHTPLLDSERKDVLREAEILH 71
          | | ::     : ||| | |  |:   || : ||:| :       :||||  :  |   |
Db     31 IDYKEIEVEEVVGRGAFGVVCKAK---WRAKDVAIKQIE-----SESERKAFIVELRQLS 82

Qy     72 KARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYP------DVAWPLRFRIL 125
          :     |  : | |  |  : :| ||   |||  :||     |       ::| |
Db     83 RVNHPNIVKLYGACLNP--VCLVMEYAEGGSLYNVLHGAEPLPYYTAAHAMSWCL----- 135

Qy    126 HEIALGVNYLHNMTP-PLLHHDLKTQNILL-DNEFHVKIADFGLSKWRMMSLSQSRSSKS 183
           : : || |||:| |  |:| |||  |:||      :||  ||| :      :: ::
Db    136 -QCSQGVAYLHSMQPKALIHRDLKPPNLLLVAGGTVLKICDFGTACDIQTHMTNNK---- 190

Qy    184 APEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNP-LQIMYSV 242
              |:  :| || :|    |  | | |::|: :| |||::|::||:::  |  :||::|
Db    191 ----GSAAWMAPEVFE---GSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWAV 243

Qy    243 SQGHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEPVLRTF----E 298
            | || :  ::||  |       ||:   |:::| :|||  :  :  :  ::| |    |
Db    244 HNGTRPPL-IKNLPKPIE------SLMTRCWSKDPSQRPSMEEIVKIMTHLMRYFPGADE 296

Qy    299 EI-------------------TFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLNIPVNH 339
           :                   :|::        |  :    :   :     |   :  |
Db    297 PLQYPCQYSDEGQSNSATSTGSFMDIASTNTSNKSDTNMEQVPATNDTIKRLESKLLKNQ 356

Qy    340 GPQEESCGSSQLHENSGSPETSRSLPAPQDNDFLSRKAQDCYFMKLHHCPGNHS------ 393
            |:   |   |  : ||  :  |||   :   :|    :    ::    ||
Db    357 AKQQSESGRLSLGASRGS--SVESLPPTSEGKRMSADMSEIE-ARIAATTGNGQPRRRSI 413

Qy    394 WDSTISGSQRAAFCDHKTTPCSSAIINPLSTAGNSERLQPGIAQQWIQSKREDIVNQMTE 453
           | |::|::        ::|    |     |:    |  |        :   |    |
Db    414 QDLTVTGTEPGQVSSRSSSPSVRMITTSGPTSEKPTRSHPWTPDDSTDTNGSDNSIPMAY 473

Qy    454 ACLNQSLD------------ALLSRDLIMKEDYELVSTKPTRTSKVRQLLDTTDIQGEEF 501
            |:  |             |:  :   | ::|  | |      ::  ||     : :|
Db    474 LTLDHQLQPLAPCPNSKESMAVFEQHCKMAQEYMKVQT------EIALLLQ----RKQEL 523

Qy    502 AKVIVQKLKD--NKQMGLQPYPEILVVSRSPSLNLLQNK 538
             : |  ||  |    :| :  ::|  ::| |    | |
Db    524 VAELDQDEKDQQNTSRLVQEHKKLLDENKSLSTYYQQCK 562


RESULT 15
T04683
hypothetical protein F8D20.290 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-May-1999
C;Accession: T04683
R;Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Jesse, T.; Heijnen, L.; Vos, P.; Mew
submitted to the Protein Sequence Database, July 1998
A;Reference number: Z15381
A;Accession: T04683
A;Molecule type: DNA
A;Residues: 1-553 <BEV>
A;Cross-references: EMBL:AL031135
A;Experimental source: cultivar Columbia; BAC clone F8D20
C;Genetics:
A;Map position: 4
A;Introns: 69/2; 107/3; 176/2; 194/3; 231/3; 289/2; 325/2; 350/1; 408/3; 440/3; 467/3
A;Note: F8D20.290


  Query Match             10.9%;  Score 307;  DB 2;  Length 553;
  Best Local Similarity   29.0%;  Pred. No. 1e-10;
  Matches   88;  Conservative   59;  Mismatches  112;  Indels   44;  Gaps   11;

Qy      7 CSALPT-------IPYHKLADLRYLSRGASGTVSSARHADWRVQVAVKHLHIHTPLLDSE 59
          |  :||       |   :|   : :: |: | :    :     :||:|       :|  |
Db    257 CVEIPTDGTDEWEIDMKQLKIEKKVACGSYGELFRGTYCS--QEVAIK-------ILKPE 307

Qy     60 R------KDVLREAEILHKARFSYIFPILGICNEPEFLGIVTEYMPNGSLNELLHRKTEY 113
          |      ::  :|  |: | |   :   :| |     | ||||:|  ||: : ||:
Db    308 RVNAEMLREFSQEVYIMRKVRHKNVVQFIGACTRSPNLCIVTEFMTRGSIYDFLHKHKGV 367

Qy    114 PDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHVKIADFGLSKWRMM 173
            :   |:  :  ::: |:||||     ::| |||| |:|:|    ||:||||:::
Db    368 FKIQSLLKVAL--DVSKGMNYLHQNN--IIHRDLKTANLLMDEHEVVKVADFGVAR---- 419

Qy    174 SLSQSRSSKSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVT 233
             |: |     | ||  :| ||  |         : |::|||:: ||:|: : |:  :|
Db    420 --VQTESGVMTAETGTYRWMAPEVIEHKPYDH---RADVFSYAIVLWELLTGELPYSYLT 474

Qy    234 NPLQIMYSVSQ-GHRPVINEESLPYDIPHRARMISLIESGWAQNPDERPSFLKCLIELEP 292
           |||    | | | || | :|: |       ::  |:|  | |:|  ||:| : :  |
Db    475 -PLQAAVGVVQKGLRPKIPKETHP-------KLTELLEKCWQQDPALRPNFAEIIEMLNQ 526
```

```
Qy     293 VLR 295
           ::|
Db     527 LIR 529


Search completed: June 14, 2001, 08:05:48
Job time: 80 sec
```

```
OM of: US-09-445-223-1 to: GenEmbl:*   out_format : pfs

Date: Jun 13, 2001  6:34 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn1_1/USPTO_spool/US09445223/runat_13062001_092033_29644/app_query.fasta_1.604
 -DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -MINLEN=0 -MAXLEN=2000000000
 -USER=US09445223_@CGN1_1_6920 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
 -WAIT -THREADS=1

Search information block:
Query: US-09-445-223-1
Query length: 540
Database: GenEmbl:*
Database sequences: 1283235
Database length: -1216004940
Search time (sec): 2258.700000


score_list:
Sequence            Strd Orig    ZScore   EScore Len   ! Documentation  ..
gb_pat1:A82777        + 2829.00 3907.62  1.9e-209  2098   ! A82777 Sequence 2 from Patent
gb_pr4:AF078530       + 2823.00 3901.83  4.1e-209  1623   ! AF078530 Homo sapiens receptor
gb_pr4:AF064824       + 2823.00 3900.27  5.0e-209  1902   ! AF064824 Homo sapiens CARD-con
gb_pr4:AF027706       + 2823.00 3897.56  7.0e-209  2501   ! AF027706 Homo sapiens serine/t
gb_pr1:AC004003       +  726.00  951.50  8.8e-45   116650  ! AC004003 Homo sapiens BAC cl
gb_pr4:AF117829       +  726.00  941.53  3.1e-44   320250  ! AF117829 Homo sapiens 8q21.3
gb_ro1:AF302127       +  577.00  779.32  3.4e-35   3559   ! AF302127 Mus musculus PKC-regu
gb_pr8:HSA278016      +  549.00  739.64  5.5e-33   3879   ! AJ278016 Homo sapiens mRNA for
gb_ro1:AF178953       +  402.00  543.18  4.9e-22   1833   ! AF178953 Mus musculus receptor
gb_pat2:I68122        +  397.00  534.14  1.5e-21   2268   ! I68122 Sequence 14 from patent
gb_ro1:MMU25995       +  397.00  534.14  1.5e-21   2268   ! U25995 Mus musculus cell death
gb_pat1:AR105328      +  395.00  535.08  1.4e-21   1557   ! AR105328 Sequence 2 from paten
gb_pat2:AX067677      +  395.00  535.08  1.4e-21   1557   ! AX067677 Sequence 2 from Paten
gb_pat1:AR105327      +  392.50  529.79  2.7e-21   1873   ! AR105327 Sequence 1 from paten
gb_pat2:AX067676      +  392.50  529.79  2.7e-21   1873   ! AX067676 Sequence 1 from Paten
gb_pl1:AF110519       +  389.50  521.78  7.6e-21   2767   ! AF110519 Lycopersicon esculent
gb_pl1:AF110518       +  389.50  521.54  7.8e-21   2836   ! AF110518 Lycopersicon esculent
gb_pl1:AF096250       +  388.50  520.17  9.3e-21   2829   ! AF096250 Lycopersicon esculent
gb_pl3:LETYRKINA      +  388.50  519.87  9.7e-21   2917   ! Y13273 Lycopersicon esculentum
gb_pr4:AF156884       +  381.00  515.67  1.7e-20   1557   ! AF156884 Homo sapiens RIP-like
gb_pat2:I14046        +  376.00  502.15  9.4e-20   3033   ! I14046 Sequence 1 from patent
gb_pat2:I35764        +  376.00  502.15  9.4e-20   3033   ! I35764 Sequence 1 from patent
gb_pl2:ATHCTR1A       +  376.00  502.15  9.4e-20   3033   ! L08789 Arabidopsis thaliana ne
gb_ro1:AF036537       +  372.00  500.65  1.1e-19   2014   ! AF036537 Rattus norvegicus hom
gb_pat2:I68123        +  369.00  495.90  2.1e-19   2137   ! I68123 Sequence 16 from patent
gb_pr9:HSU50062       +  369.00  493.90  2.7e-19   2617   ! U50062 Homo sapiens RIP protei
gb_pat1:AR044139      +  367.00  516.03  1.6e-20   210   ! AR044139 Sequence 31 from paten
gb_pl2:AF305913       +  359.00  478.42  2.0e-18   3081   ! AF305913 Arabidopsis thaliana
gb_pl3:LEAJ5077       +  345.50  458.56  2.5e-17   3456   ! AJ005077 Lycopersicon esculent
gb_pl2:AF305912       +  341.50  453.36  4.9e-17   3338   ! AF305912 Hordeum vulgare EDR1
gb_pr4:AF116826       +  337.50  448.86  8.7e-17   3001   ! AF116826 Homo sapiens clone HH
gb_pat1:AX056437      +  334.50  446.47  1.2e-16   2508   ! AX056437 Sequence 81 from Pate
gb_pl2:AF305911       +  330.50  438.65  3.2e-16   3158   ! AF305911 Oryza sativa EDR1 (ED
gb_pl2:ATY14199       +  328.50  443.09  1.8e-16   1521   ! Y14199 Arabidopsis thaliana mR
gb_pr1:AB001872       +  328.50  434.67  5.4e-16   3569   ! AB001872 Homo sapiens mRNA for
gb_in2:AF199466       +  327.00  433.22  6.5e-16   3349   ! AF199466 Drosophila melanogast
gb_pl3:SOYPROKIN      +  318.00  427.29  1.4e-15   1724   ! M67449 Glycine max protein kin
gb_pl1:AC004697       -  316.50  384.48  3.4e-13   106718  ! AC004697 Arabidopsis thalian
gb_pl3:FSY298992      +  315.50  424.50  2.0e-15   1610   ! AJ298992 Fagus sylvatica parti
gb_in3:DDIDPYK1A      +  313.50  425.58  1.7e-15   1090   ! M33785 D.discoideum protein-ty
```

```
gb_ro1:RATPK          +  313.50  413.37  8.3e-15   3754   ! D49785 Rattus norvegicus mF
gb_in3:DDU32174       +  313.50  408.85  1.5e-14   5931   ! U32174 Dictyostelium discoi
gb_pr1:AB009356       +  312.00  414.29  7.3e-15   2769   ! AB009356 Homo sapiens mRNA
gb_pr8:HSAJ5282       +  311.50  411.34  1.1e-14   3482   ! AJ005282 Homo sapiens mRNA
gb_pr9:HUMGUANCYC     +  311.50  409.77  1.3e-14   4081   ! L13436 Homo sapiens guanyla

seq_name: gb_pat1:A82777

seq_documentation_block:
LOCUS       A82777        2098 bp    DNA             PAT       21-JAN-2000
DEFINITION  Sequence 2 from Patent WO9855507.
ACCESSION   A82777
VERSION     A82777.1  GI:6732464
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2098)
  AUTHORS   Boldin,M. and Wallach,D.
  TITLE     MODULATORS OF INTRACELLULAR INFLAMMATION, CELL DEATH AND CELL
            SURVIVAL PATHWAYS
  JOURNAL   Patent: WO 9855507-A 2 10-DEC-1998;
            BOLDIN MARK (IL); WALLACH DAVID (IL)
FEATURES             Location/Qualifiers
     source          1. .2098
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      649 a    452 c    449 g    539 t      9 others
ORIGIN


alignment_scores:
            Quality: 2829.00              Length:    540
              Ratio:  5.239                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-445-223-1 x A82777   ..

 Align seg 1/1  to: A82777  from: 1  to: 2098

       1 MetAsnGlyGluAlaIleCysSerAlaLeuProThrIleProTyrHisLy 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     260 ATGAACGGGGAGGCCATCTGCAGCGCCCTGCCCACCATTCCCTACCACAA 309

      17 sLeuAlaAspLeuArgTyrLeuSerArgGlyAlaSerGlyThrValSerS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     310 ACTCGCCGACCTGCGCTACCTGAGCCGCGGCGCCTCTGGCACTGTGTCGT 359

      34 erAlaArgHisAlaAspTrpArgValGlnValAlaValLysHisLeuHis 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     360 CCGCCCGCCACGCAGACTGGCGCGTCCAGGTGGCCGTGAAGCACCTGCAC 409

      51 IleHisThrProLeuLeuAspSerGluArgLysAspValLeuArgGluAl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     410 ATCCACACTCCGCTGCTCGACAGTGAAAGAAAGGATGTTTTAAGAGAAGC 459

      67 aGluIleLeuHisLysAlaArgPheSerTyrIlePheProIleLeuGlyI 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     460 TGAAATTTTACACAAAGCTAGATTTAGTTACATTTTTCCAATTTTGGGAA 509

      84 leCysAsnGluProGluPheLeuGlyIleValThrGluTyrMetProAsn 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     510 TTTGCAATGAGCCTGAATTTTTGGGAATAGTTACTGAATACATGCCAAAT 559

     101 GlySerLeuAsnGluLeuLeuHisArgLysThrGluTyrProAspValAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     560 GGATCATTAAATGAACTCCTACATAGGAAAACTGAATATCCTGATGTTGC 609

     117 aTrpProLeuArgPheArgIleLeuHisGluIleAlaLeuGlyValAsnT 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     610 TTGGCCATTGAGATTTCGCATCCTGCATGAAATTGCCCTTGGTGTAAATT 659
```

```
 134 yrLeuHisAsnMetThrProProLeuLeuHisHisAspLeuLysThrGln 150
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 660 ACCTGCACAATATGACTCCTCCTTTACTTCATCATGACTTGAAGACTCAG 709

 151 AsnIleLeuLeuAspAsnGluPheHisValLysIleAlaAspPheGlyLe 167
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 710 AATATCTTATTGGACAATGAATTTCATGTTAAGATTGCAGATTTTGGTTT 759

 167 uSerLysTrpArgMetMetSerLeuSerGlnSerArgSerSerLysSerA 184
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 760 ATCAAAGTGGCGCATGATGTCCCTCTCACAGTCACGAAGTAGCAAATCTG 809

 184 laProGluGlyGlyThrIleIleTyrMetProProGluAsnTyrGluPro 200
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 810 CACCAGAAGGAGGGACAATTATTTATATGCCACCTGAAAACTATGAACCT 859

 201 GlyGlnLysSerArgAlaSerIleLysHisAspIleTyrSerTyrAlaVa 217
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 860 GGACAAAAATCAAGGGCCAGTATCAAGCACGATATATATAGCTATGCAGT 909

 217 lIleThrTrpGluValLeuSerArgLysGlnProPheGluAspValThrA 234
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 910 TATCACATGGGAAGTGTTATCCAGAAAACAGCCTTTTGAAGATGTCACCA 959

 234 snProLeuGlnIleMetTyrSerValSerGlnGlyHisArgProValIle 250
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 960 ATCCTTTGCAGATAATGTATAGTGTGTCACAAGGACATCGACCTGTTATT 1009

 251 AsnGluGluSerLeuProTyrAspIleProHisArgAlaArgMetIleSe 267
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1010 AATGAAGAAAGTTTGCCATATGATATACCTCACCGAGCACGTATGATCTC 1059

 267 rLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerPheL 284
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1060 TCTAATAGAAAGTGGATGGGCACAAAATCCAGATGAAAGACCATCTTTCT 1109

 284 euLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGluIle 300
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1110 TAAAATGTTTAATAGAACTTGAACCAGTTTTGAGAACATTTGAAGAGATA 1159

 301 ThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSerVa 317
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1160 ACTTTTCTTGAAGCTGTTATTCAGCTAAAGAAAACAAAGTTACAGAGTGT 1209

 317 lSerSerAlaIleHisLeuCysAspLysLysLysMetGluLeuSerLeuA 334
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1210 TTCAAGTGCCATTCACCTATGTGACAAGAAGAAAATGGAATTATCTCTGA 1259

 334 snIleProValAsnHisGlyProGlnGluGluSerCysGlySerSerGln 350
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1260 ACATACCTGTAAATCATGGTCCACAAGAGGAATCATGTGGATCCTCTCAG 1309

 351 LeuHisGluAsnSerGlySerProGluThrSerArgSerLeuProAlaPr 367
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1310 CTCCATGAAAATAGTGGTTCTCCTGAAACTTCAAGGTCCCTGCCAGCTCC 1359

 367 oGlnAspAsnAspPheLeuSerArgLysAlaGlnAspCysTyrPheMetL 384
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1360 TCAAGACAATGATTTTTTATCTAGAAAAGCTCAAGACTGTTATTTTATGA 1409

 384 ysLeuHisHisCysProGlyAsnHisSerTrpAspSerThrIleSerGly 400
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1410 AGCTGCATCACTGTCCTGGAAATCACAGTTGGGATAGCACCATTTCTGGA 1459

 401 SerGlnArgAlaAlaPheCysAspHisLysThrThrProCysSerSerAl 417
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1460 TCTCAAAGGGCTGCATTCTGTGATCACAAGACCACTCCATGCTCTTCAGC 1509

 417 aIleIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeuGlnProG 434
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1510 AATAATAAATCCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAGCCTG 1559

 434 lyIleAlaGlnGlnTrpIleGlnSerLysArgGluAspIleValAsnGln 450
```

```
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1560 GTATAGCCCAGCAGTGGATCCAGAGCAAAAGGGAAGACATTGTGAACCAA 1609

 451 MetThrGluAlaCysLeuAsnGlnSerLeuAspAlaLeuLeuSerArgAs 467
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1610 ATGACAGAAGCCTGCCTTAACCAGTCGCTAGATGCCCTTCTGTCCAGGGA 1659

 467 pLeuIleMetLysGluAspTyrGluLeuValSerThrLysProThrArgT 484
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1660 CTTGATCATGAAAGAGGACTATGAACTTGTTAGTACCAAGCCTACAAGGA 1709

 484 hrSerLysValArgGlnLeuLeuAspThrThrAspIleGlnGlyGluGlu 500
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1710 CCTCAAAAGTCAGACAATTACTAGACACTACTGACATCCAAGGAGAAGAA 1759

 501 PheAlaLysValIleValGlnLysLeuLysAspAsnLysGlnMetGlyLe 517
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1760 TTTGCCAAAGTTATAGTACAAAAATTGAAAGATAACAAACAAATGGGTCT 1809

 517 uGlnProTyrProGluIleLeuValValSerArgSerProSerLeuAsnL 534
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1810 TCAGCCTTACCCGGAAATACTTGTGGTTTCTAGATCACCATCTTTAAATT 1859

 534 euLeuGlnAsnLysSerMet 540
     ||||||||||||||||||||
1860 TACTTCAAAATAAAAGCATG 1879
```

```
seq_name: gb_pr4:AF078530

seq_documentation_block:
LOCUS       AF078530     1623 bp    mRNA            PRI       28-JUL-1998
DEFINITION  Homo sapiens receptor interacting protein 2 (RIP2) mRNA, complete
            cds.
ACCESSION   AF078530
VERSION     AF078530.1  GI:3342909
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1623)
  AUTHORS   McCarthy,J.V., Ni,J. and Dixit,V.M.
  TITLE     RIP2 is a novel NF-kappaB-activating and cell death-inducing kinase
  JOURNAL   J. Biol. Chem. 273 (27), 16968-16975 (1998)
  MEDLINE   98307936
REFERENCE   2  (bases 1 to 1623)
  AUTHORS   McCarthy,J.V., Ni,J. and Dixit,V.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUL-1998) Molecular Oncology, Genentech Inc, 1 DNA
            Way, South San Francisco, CA 94080, USA
FEATURES             Location/Qualifiers
     source          1..1623
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            1..1623
                     /gene="RIP2"
     CDS             1..1623
                     /gene="RIP2"
                     /function="activtes NF-kappaB"
                     /function="induces cell death"
                     /note="RIP2; serine/threonine kinase homolog; contains
                     CARD motif"
                     /codon_start=1
                     /product="receptor interacting protein 2"
                     /protein_id="AAC27722.1"
                     /db_xref="GI:3342910"
                     /translation="MNGEAICSALPTIPYHKLADLRYLSRGASGTVSSARHADWRVQV
                     AVKHLHIHTPLLDSERKDVLREAEILHKARFSYILPILGICNEPEFLGIVTEYMPNGS
                     LNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHV
                     KIADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVI
                     TWEVLSRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIPHRARMISLIESGWAQN
                     PDERPSFLKCLIELEPVLRTFEEITFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLN
                     IPVNHGPQEESCGSSQLHENSGSPETSRSLPAPQDNDFLSRKAQDCYFMKLHHCPGNH
```

```
                    SWDSTISGSQRAAFCDHKTTPCSSAIINPLSTAGNSERLQPGIAQQWIQSKREDIVNQ
                    MTEACLNQSLDALLSRDLIMKEDYELVSTKPTRTSKVRQLLDTTDIQGEEFAKVIVQK
                    LKDNKQMGLQPYPEILVVSRSPSLNLLQNKSM"
BASE COUNT      519 a    362 c    320 g    422 t
ORIGIN


alignment_scores:
           Quality: 2823.00              Length:    540
             Ratio:  5.228                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.815

alignment_block:
 US-09-445-223-1 x AF078530   ..

 Align seg 1/1  to: AF078530  from: 1  to: 1623

      1 MetAsnGlyGluAlaIleCysSerAlaLeuProThrIleProTyrHisLy 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
      1 ATGAACGGGGAGGCCATCTGCAGCGCCCTGCCCACCATTCCCTACCACAA 50

     17 sLeuAlaAspLeuArgTyrLeuSerArgGlyAlaSerGlyThrValSerS 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     51 ACTCGCCGACCTGCGCTACCTGAGCCGCGGCGCCTCTGGCACTGTGTCGT 100

     34 erAlaArgHisAlaAspTrpArgValGlnValAlaValLysHisLeuHis 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    101 CCGCCCGCCACGCAGACTGGCGCGTCCAGGTGGCCGTGAAGCACCTGCAC 150

     51 IleHisThrProLeuLeuAspSerGluArgLysAspValLeuArgGluAl 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    151 ATCCACACTCCGCTGCTCGACAGTGAAAGAAAGGATGTCTTAAGAGAAGC 200

     67 aGluIleLeuHisLysAlaArgPheSerTyrIlePheProIleLeuGlyI 84
        ||||||||||||||||||||||||||||||||||:::|||||||||||||
    201 TGAAATTTACACAAAGCTAGATTTAGTTACATTCTTCCAATTTTGGGAA 250

     84 leCysAsnGluProGluPheLeuGlyIleValThrGluTyrMetProAsn 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    251 TTTGCAATGAGCCTGAATTTTTGGGAATAGTTACTGAATACATGCCAAAT 300

    101 GlySerLeuAsnGluLeuLeuHisArgLysThrGluTyrProAspValAl 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    301 GGATCATTAAATGAACTCCTACATAGGAAAACTGAATATCCTGATGTTGC 350

    117 aTrpProLeuArgPheArgIleLeuHisGluIleAlaLeuGlyValAsnT 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    351 TTGGCCATTGAGATTTCGCATCCTGCATGAAATTGCCCTTGGTGTAAATT 400

    134 yrLeuHisAsnMetThrProProLeuLeuHisHisAspLeuLysThrGln 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    401 ACCTGCACAATATGACTCCTCCTTTACTTCATCATGACTTGAAGACTCAG 450

    151 AsnIleLeuLeuAspAsnGluPheHisValLysIleAlaAspPheGlyLe 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    451 AATATCTTATTGGACAATGAATTTCATGTTAAGATTGCAGATTTTGGTTT 500

    167 uSerLysTrpArgMetMetSerLeuSerGlnSerArgSerSerLysSerA 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    501 ATCAAAGTGGCGCATGATGTCCCTCTCACAGTCACGAAGTAGCAAATCTG 550

    184 laProGluGlyGlyThrIleIleTyrMetProProGluAsnTyrGluPro 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    551 CACCAGAAGGAGGGACAATTATCTATATGCCACCTGAAAACTATGAACCT 600

    201 GlyGlnLysSerArgAlaSerIleLysHisAspIleTyrSerTyrAlaVa 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    601 GGACAAAAATCAAGGGCCAGTATCAAGCACGATATATATAGCTATGCAGT 650

    217 lIleThrTrpGluValLeuSerArgLysGlnProPheGluAspValThrA 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    651 TATCACATGGGAAGTGTTATCCAGAAAACAGCCTTTTGAAGATGTCACCA 700

    234 snProLeuGlnIleMetTyrSerValSerGlnGlyHisArgProValIle 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    701 ATCCTTTGCAGATAATGTATAGTGTGTCACAAGGACATCGACCTGTTATT 750

    251 AsnGluGluSerLeuProTyrAspIleProHisArgAlaArgMetIleSe 267
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    751 AATGAAGAAAGTTTGCCATATGATATACCTCACCGAGCACGTATGATCTC 800

    267 rLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerPheL 284
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    801 TCTAATAGAAAGTGGATGGGCACAAAATCCAGATGAAAGACCATCTTTCT 850

    284 euLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGluIle 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    851 TAAAATGTTTAATAGAACTTGAACCAGTTTTGAGAACATTTGAAGAGATA 900

    301 ThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSerVa 317
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    901 ACTTTTCTTGAAGCTGTTATTCAGCTAAAGAAAACAAAGTTACAGAGTGT 950

    317 lSerSerAlaIleHisLeuCysAspLysLysLysMetGluLeuSerLeuA 334
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    951 TTCAAGTGCCATTCACCTATGTGACAAGAAGAAAATGGAATTATCTCTGA 1000

    334 snIleProValAsnHisGlyProGlnGluGluSerCysGlySerSerGln 350
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1001 ACATACCTGTAAATCATGGTCCACAAGAGGAATCATGTGGATCCTCTCAG 1050

    351 LeuHisGluAsnSerGlySerProGluThrSerArgSerLeuProAlaPr 367
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1051 CTCCATGAAAATAGTGGTTCTCCTGAAACTTCAAGGTCCCTGCCAGCTCC 1100

    367 oGlnAspAsnAspPheLeuSerArgLysAlaGlnAspCysTyrPheMetL 384
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1101 TCAAGACAATGATTTTTTATCTAGAAAAGCTCAAGACTGTTATTTTATGA 1150

    384 ysLeuHisHisCysProGlyAsnHisSerTrpAspSerThrIleSerGly 400
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1151 AGCTGCATCACTGTCCTGGAAATCACAGTTGGGACAGCACCATTTCTGGA 1200

    401 SerGlnArgAlaAlaPheCysAspHisLysThrThrProCysSerSerAl 417
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1201 TCTCAAAGGGCTGCATTCTGTGATCACAAGACCACTCCATGCTCTTCAGC 1250

    417 aIleIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeuGlnProG 434
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1251 AATAATAAATCCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAGCCTG 1300

    434 lyIleAlaGlnGlnTrpIleGlnSerLysArgGluAspIleValAsnGln 450
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1301 GTATAGCCCAGCAGTGGATCCAGAGCAAAAGGGAAGACATTGTGAACCAA 1350

    451 MetThrGluAlaCysLeuAsnGlnSerLeuAspAlaLeuLeuSerArgAs 467
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1351 ATGACAGAAGCCTGCCTTAACCAGTCGCTAGATGCCCTTCTGTCCAGGGA 1400

    467 pLeuIleMetLysGluAspTyrGluLeuValSerThrLysProThrArgT 484
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1401 CTTGATCATGAAAGAGGACTATGAACTTGTTAGTACCAAGCCTACAAGGA 1450

    484 hrSerLysValArgGlnLeuLeuAspThrThrAspIleGlnGlyGluGlu 500
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1451 CCTCAAAAGTCAGACAATTACTAGACACTACTGACATCCAAGGAGAAGAA 1500

    501 PheAlaLysValIleValGlnLysLeuLysAspAsnLysGlnMetGlyLe 517
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1501 TTTGCCAAAGTTATAGTACAAAAATTGAAAGATAACAAACAAATGGGTCT 1550

    517 uGlnProTyrProGluIleLeuValValSerArgSerProSerLeuAsnL 534
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1551 TCAGCCTTACCCGGAAATACTTGTGGTTTCTAGATCACCATCTTTAAATT 1600

    534 euLeuGlnAsnLysSerMet 540
```

```
           ||||||||||||||||||||
      1601 TACTTCAAAATAAAAGCATG 1620

seq_name: gb_pr4:AF064824

seq_documentation_block:
LOCUS       AF064824     1902 bp    mRNA            PRI       07-JUL-1998
DEFINITION  Homo sapiens CARD-containing ICE associated kinase mRNA, complete
            cds.
ACCESSION   AF064824
VERSION     AF064824.1  GI:3290171
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1902)
  AUTHORS   Thome,M., Hofmann,K., Burns,K., Martinon,F., Bodmer,J.-L.,
            Mattmann,C. and Tschopp,J.
  TITLE     Identification of CARDIAK, a RIP-like kinase that associates with
            caspase-1
  JOURNAL   Curr. Biol. 8 (1998) In press
REFERENCE   2  (bases 1 to 1902)
  AUTHORS   Thome,M., Hofmann,K., Burns,K., Martinon,F., Bodmer,J.-L.,
            Mattmann,C. and Tschopp,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAY-1998) Institute of Biochemistry, University of
            Lausanne, Chemin des Boveresses, 155, Epalinges, CH 1066,
            Switzerland
FEATURES             Location/Qualifiers
     source          1. .1902
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /db_xref="dbEST:AA161113"
                     /clone_lib="Stratagene library, catalog number 937208"
                     /cell_type="pancreatic adenocarcinoma"
     CDS             243. .1865
                     /note="CARDIAK; RIP-like kinase"
                     /codon_start=1
                     /product="CARD-containing ICE associated kinase"
                     /protein_id="AAC25668.1"
                     /db_xref="GI:3290172"
                     /translation="MNGEAICSALPTIPYHKLADLRYLSRGASGTVSSARHADWRVQV
                     AVKHLHIHTPLLDSERKDVLREAEILHKARFSYILPILGICNEPEFLGIVTEYMPNGS
                     LNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHV
                     KIADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVI
                     TWEVLSRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIPHRARMISLIESGWAQN
                     PDERPSFLKCLIELEPVLRTFEEITFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLN
                     IPVNHGPQEESCGSSQLHENSGSPETSRSLPAPQDNDFLSRKAQDCYFMKLHHCPGNH
                     SWDSTISGSQRAAFCDHKTTPCSSAIINPLSTAGNSERLQPGIAQQWIQSKREDIVNQ
                     MTEACLNQSLDALLSRDLIMKEDYELVSTKPTRTSKVRQLLDTTDIQGEEFAKVIVQK
                     LKDNKQMGLQPYPEILVVSRSPSLNLLQNKSM"
     misc_feature    294. .1124
                     /note="encodes kinase domain"
     misc_feature    1125. .1544
                     /note="encodes intermediate domain"
     misc_feature    1545. .1811
                     /note="encodes CARD motif"
BASE COUNT      582 a    432 c    419 g    468 t      1 others
ORIGIN


alignment_scores:
           Quality: 2823.00              Length:    540
             Ratio:  5.228                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.815

alignment_block:
 US-09-445-223-1 x AF064824   ..

 Align seg 1/1  to: AF064824  from: 1  to: 1902

      1 MetAsnGlyGluAlaIleCysSerAlaLeuProThrIleProTyrHisLy 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    243 ATGAACGGGGAGGCCATCTGCAGCGCCCTGCCCACCATTCCCTACCACAA 292

     17 sLeuAlaAspLeuArgTyrLeuSerArgGlyAlaSerGlyThrValSerS 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    293 ACTCGCCGACCTGCGCTACCTGAGCCGCGGCGCCTCTGGCACTGTGTCGT 342

     34 erAlaArgHisAlaAspTrpArgValGlnValAlaValLysHisLeuHis 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    343 CCGCCCGCCACGCAGACTGGCGCGTCCAGGTGGCCGTGAAGCACCTGCAC 392

     51 IleHisThrProLeuLeuAspSerGluArgLysAspValLeuArgGluAl 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    393 ATCCACACTCCGCTGCTCGACAGTGAAAGAAAGGATGTCTTAAGAGAAGC 442

     67 aGluIleLeuHisLysAlaArgPheSerTyrIlePheProIleLeuGlyI 84
        ||||||||||||||||||||||||||||||||||:::|||||||||||||
    443 TGAAATTTACACAAAGCTAGATTTAGTTACATTCTTCCAATTTTGGGAA 492

     84 leCysAsnGluProGluPheLeuGlyIleValThrGluTyrMetProAsn 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    493 TTTGCAATGAGCCTGAATTTTTGGGAATAGTTACTGAATACATGCCAAAT 542

    101 GlySerLeuAsnGluLeuLeuHisArgLysThrGluTyrProAspValAl 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    543 GGATCATTAAATGAACTCCTACATAGGAAAACTGAATATCCTGATGTTGC 592

    117 aTrpProLeuArgPheArgIleLeuHisGluIleAlaLeuGlyValAsnT 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    593 TTGGCCATTGAGATTTCGCATCCTGCATGAAATTGCCCTTGGTGTAAATT 642

    134 yrLeuHisAsnMetThrProProLeuLeuHisHisAspLeuLysThrGln 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    643 ACCTGCACAATATGACTCCTCCTTTACTTCATCATGACTTGAAGACTCAG 692

    151 AsnIleLeuLeuAspAsnGluPheHisValLysIleAlaAspPheGlyLe 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    693 AATATCTTATTGGACAATGAATTTCATGTTAAGATTGCAGATTTTGGTTT 742

    167 uSerLysTrpArgMetMetSerLeuSerGlnSerArgSerSerLysSerA 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    743 ATCAAAGTGGCGCATGATGTCCCTCTCACAGTCACGAAGTAGCAAATCTG 792

    184 laProGluGlyGlyThrIleIleTyrMetProProGluAsnTyrGluPro 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    793 CACCAGAAGGAGGGACAATTATCTATATGCCACCTGAAAACTATGAACCT 842

    201 GlyGlnLysSerArgAlaSerIleLysHisAspIleTyrSerTyrAlaVa 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    843 GGACAAAAATCAAGGGCCAGTATCAAGCACGATATATATAGCTATGCAGT 892

    217 lIleThrTrpGluValLeuSerArgLysGlnProPheGluAspValThrA 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    893 TATCACATGGGAAGTGTTATCCAGAAAACAGCCTTTTGAAGATGTCACCA 942

    234 snProLeuGlnIleMetTyrSerValSerGlnGlyHisArgProValIle 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    943 ATCCTTTGCAGATAATGTATAGTGTGTCACAAGGACATCGACCTGTTATT 992

    251 AsnGluGluSerLeuProTyrAspIleProHisArgAlaArgMetIleSe 267
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    993 AATGAAGAAAGTTTGCCATATGATATACCTCACCGAGCACGTATGATCTC 1042

    267 rLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerPheL 284
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1043 TCTAATAGAAAGTGGATGGGCACAAAATCCAGATGAAAGACCATCTTTCT 1092

    284 euLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGluIle 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1093 TAAAATGTTTAATAGAACTTGAACCAGTTTTGAGAACATTTGAAGAGATA 1142

    301 ThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSerVa 317
        ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
 1143 ACTTTTCTTGAAGCTGTTATTCAGCTAAAGAAAACAAAGTTACAGAGTGT 1192

  317 lSerSerAlaIleHisLeuCysAspLysLysLysMetGluLeuSerLeuA 334
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1193 TTCAAGTGCCATTCACCTATGTGACAAGAAGAAAATGGAATTATCTCTGA 1242

  334 snIleProValAsnHisGlyProGlnGluGluSerCysGlySerSerGln 350
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1243 ACATACCTGTAAATCATGGTCCACAAGAGGAATCATGTGGATCCTCTCAG 1292

  351 LeuHisGluAsnSerGlySerProGluThrSerArgSerLeuProAlaPr 367
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1293 CTCCATGAAAATAGTGGTTCTCCTGAAACTTCAAGGTCCCTGCCAGCTCC 1342

  367 oGlnAspAsnAspPheLeuSerArgLysAlaGlnAspCysTyrPheMetL 384
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1343 TCAAGACAATGATTTTTTATCTAGAAAAGCTCAAGACTGTTATTTTATGA 1392

  384 ysLeuHisHisCysProGlyAsnHisSerTrpAspSerThrIleSerGly 400
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1393 AGCTGCATCACTGTCCTGGAAATCACAGTTGGGATAGCACCATTTCTGGA 1442

  401 SerGlnArgAlaAlaPheCysAspHisLysThrThrProCysSerSerAl 417
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1443 TCTCAAAGGGCTGCATTCTGTGATCACAAGACCACTCCATGCTCTTCAGC 1492

  417 aIleIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeuGlnProG 434
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1493 AATAATAAATCCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAGCCTG 1542

  434 lyIleAlaGlnGlnTrpIleGlnSerLysArgGluAspIleValAsnGln 450
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1543 GTATAGCCCAGCAGTGGATCCAGAGCAAAAGGGAAGACATTGTGAACCAA 1592

  451 MetThrGluAlaCysLeuAsnGlnSerLeuAspAlaLeuLeuSerArgAs 467
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1593 ATGACAGAAGCCTGCCTTAACCAGTCGCTAGATGCCCTTCTGTCCAGGGA 1642

  467 pLeuIleMetLysGluAspTyrGluLeuValSerThrLysProThrArgT 484
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1643 CTTGATCATGAAAGAGGACTATGAACTTGTTAGTACCAAGCCTACAAGGA 1692

  484 hrSerLysValArgGlnLeuLeuAspThrThrAspIleGlnGlyGluGlu 500
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1693 CCTCAAAAGTCAGACAATTACTAGACACTACTGACATCCAAGGAGAAGAA 1742

  501 PheAlaLysValIleValGlnLysLeuLysAspAsnLysGlnMetGlyLe 517
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1743 TTTGCCAAAGTTATAGTACAAAAATTGAAAGATAACAAACAAATGGGTCT 1792

  517 uGlnProTyrProGluIleLeuValValSerArgSerProSerLeuAsnL 534
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1793 TCAGCCTTACCCGGAAATACTTGTGGTTTCTAGATCACCATCTTTAAATT 1842

  534 euLeuGlnAsnLysSerMet 540
      ||||||||||||||||||||
 1843 TACTTCAAAATAAAAGCATG 1862

seq_name: gb_pr4:AF027706

seq_documentation_block:
LOCUS       AF027706     2501 bp    mRNA            PRI       10-SEP-1998
DEFINITION  Homo sapiens serine/threonine kinase RICK (RICK) mRNA, complete
            cds.
ACCESSION   AF027706
VERSION     AF027706.1  GI:3123886
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2501)
  AUTHORS   Inohara,N., del Peso,L., Koseki,T., Chen,S. and Nunez,G.
  TITLE     RICK, a novel protein kinase containing a caspase recruitment
            domain, interacts with CLARP and regulates CD95-mediated apoptosis
  JOURNAL   J. Biol. Chem. 273 (20), 12296-12300 (1998)
  MEDLINE   98241596
REFERENCE   2  (bases 1 to 2501)
  AUTHORS   Inohara,N., Koseki,T., Chen,S., del Peso,L. and Nunez,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-OCT-1997) Dept. Pathlogy, Comprehensive Cancer &
            Geriatrics Center, 4-131 CCGC 1500 E. Medical Center Dr, Ann Arbor,
            MI 48109, USA
FEATURES             Location/Qualifiers
     source          1. .2501
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8q21"
     gene            1. .2501
                     /gene="RICK"
     CDS             225. .1847
                     /gene="RICK"
                     /function="regulates apoptosis"
                     /note="CARD domain protein; phosphorylated protein"
                     /codon_start=1
                     /product="serine/threonine kinase RICK"
                     /protein_id="AAC34970.1"
                     /db_xref="GI:3123887"
                     /translation="MNGEAICSALPTIPYHKLADLRYLSRGASGTVSSARHADWRVQV
                     AVKHLHIHTPLLDSERKDVLREAEILHKARFSYILPILGICNEPEFLGIVTEYMPNGS
                     LNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHV
                     KIADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVI
                     TWEVLSRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIPHRARMISLIESGWAQN
                     PDERPSFLKCLIELEPVLRTFEEITFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLN
                     IPVNHGPQEESCGSSQLHENSGSPETSRSLPAPQDNDFLSRKAQDCYFMKLHHCPGNH
                     SWDSTISGSQRAAFCDHKTTPCSSAIINPLSTAGNSERLQPGIAQQWIQSKREDIVNQ
                     MTEACLNQSLDALLSRDLIMKEDYELVSTKPTRTSKVRQLLDTTDIQGEEFAKVIVQK
                     LKDNKQMGLQPYPEILVVSRSPSLNLLQNKSM"
     polyA_signal    1833. .1838
                     /gene="RICK"
     polyA_site      1869
                     /gene="RICK"
     polyA_signal    2479. .2484
                     /gene="RICK"
     polyA_site      2501
                     /gene="RICK"
BASE COUNT      768 a    535 c    499 g    699 t
ORIGIN


alignment_scores:
           Quality: 2823.00              Length:    540
             Ratio:  5.228                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.815

alignment_block:
 US-09-445-223-1 x AF027706   ..

 Align seg 1/1  to: AF027706  from: 1  to: 2501

        1 MetAsnGlyGluAlaIleCysSerAlaLeuProThrIleProTyrHisLy 17
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      225 ATGAACGGGGAGGCCATCTGCAGCGCCCTGCCCACCATTCCCTACCACAA 274

       17 sLeuAlaAspLeuArgTyrLeuSerArgGlyAlaSerGlyThrValSerS 34
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      275 ACTCGCCGACCTGCGCTACCTGAGCCGCGGCGCCTCTGGCACTGTGTCGT 324

       34 erAlaArgHisAlaAspTrpArgValGlnValAlaValLysHisLeuHis 50
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      325 CCGCCCGCCACGCAGACTGGCGCGTCCAGGTGGCCGTGAAGCACCTGCAC 374

       51 IleHisThrProLeuLeuAspSerGluArgLysAspValLeuArgGluAl 67
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      375 ATCCACACTCCGCTGCTCGACAGTGAAAGAAAGGATGTCTTAAGAGAAGC 424
```

```
  67 aGluIleLeuHisLysAlaArgPheSerTyrIlePheProIleLeuGlyI 84
     ||||||||||||||||||||||||||||||||||:::|||||||||||||
 425 TGAAATTTTACACAAAGCTAGATTTAGTTACATTCTTCCAATTTTGGGAA 474

  84 leCysAsnGluProGluPheLeuGlyIleValThrGluTyrMetProAsn 100
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 475 TTTGCAATGAGCCTGAATTTTTGGGAATAGTTACTGAATACATGCCAAAT 524

 101 GlySerLeuAsnGluLeuLeuHisArgLysThrGluTyrProAspValAl 117
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 525 GGATCATTAAATGAACTCCTACATAGGAAAACTGAATATCCTGATGTTGC 574

 117 aTrpProLeuArgPheArgIleLeuHisGluIleAlaLeuGlyValAsnT 134
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 575 TTGGCCATTGAGATTTCGCATCCTGCATGAAATTGCCCTTGGTGTAAATT 624

 134 yrLeuHisAsnMetThrProProLeuLeuHisHisAspLeuLysThrGln 150
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 625 ACCTGCACAATATGACTCCTCCTTTACTTCATCATGACTTGAAGACTCAG 674

 151 AsnIleLeuLeuAspAsnGluPheHisValLysIleAlaAspPheGlyLe 167
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 675 AATATCTTATTGGACAATGAATTTCATGTTAAGATTGCAGATTTTGGTTT 724

 167 uSerLysTrpArgMetMetSerLeuSerGlnSerArgSerSerLysSerA 184
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 725 ATCAAAGTGGCGCATGATGTCCCTCTCACAGTCACGAAGTAGCAAATCTG 774

 184 laProGluGlyGlyThrIleIleTyrMetProProGluAsnTyrGluPro 200
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 775 CACCAGAAGGAGGGACAATTATCTATATGCCACCTGAAAACTATGAACCT 824

 201 GlyGlnLysSerArgAlaSerIleLysHisAspIleTyrSerTyrAlaVa 217
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 825 GGACAAAAATCAAGGGCCAGTATCAAGCACGATATATATAGCTATGCAGT 874

 217 lIleThrTrpGluValLeuSerArgLysGlnProPheGluAspValThrA 234
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 875 TATCACATGGGAAGTGTTATCCAGAAAACAGCCTTTTGAAGATGTCACCA 924

 234 snProLeuGlnIleMetTyrSerValSerGlnGlyHisArgProValIle 250
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 925 ATCCTTTGCAGATAATGTATAGTGTGTCACAAGGACATCGACCTGTTATT 974

 251 AsnGluGluSerLeuProTyrAspIleProHisArgAlaArgMetIleSe 267
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 975 AATGAAGAAAGTTTGCCATATGATATACCTCACCGAGCACGTATGATCTC 1024

 267 rLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerPheL 284
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1025 TCTAATAGAAAGTGGATGGGCACAAAATCCAGATGAAAGACCATCTTTCT 1074

 284 euLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGluIle 300
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1075 TAAAATGTTTAATAGAACTTGAACCAGTTTTGAGAACATTTGAAGAGATA 1124

 301 ThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSerVa 317
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1125 ACTTTTCTTGAAGCTGTTATTCAGCTAAAGAAAACAAAGTTACAGAGTGT 1174

 317 lSerSerAlaIleHisLeuCysAspLysLysLysMetGluLeuSerLeuA 334
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1175 TTCAAGTGCCATTCACCTATGTGACAAGAAGAAAATGGAATTATCTCTGA 1224

 334 snIleProValAsnHisGlyProGlnGluGluSerCysGlySerSerGln 350
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1225 ACATACCTGTAAATCATGGTCCACAAGAGGAATCATGTGGATCCTCTCAG 1274

 351 LeuHisGluAsnSerGlySerProGluThrSerArgSerLeuProAlaPr 367
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1275 CTCCATGAAAATAGTGGTTCTCCTGAAACTTCAAGGTCCCTGCCAGCTCC 1324

 367 oGlnAspAsnAspPheLeuSerArgLysAlaGlnAspCysTyrPheMetL 384
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1325 TCAAGACAATGATTTTTTATCTAGAAAAGCTCAAGACTGTTATTTTATGA 1374

 384 ysLeuHisHisCysProGlyAsnHisSerTrpAspSerThrIleSerGly 400
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1375 AGCTGCATCACTGTCCTGGAAATCACAGTTGGGATAGCACCATTTCTGGT 1424

 401 SerGlnArgAlaAlaPheCysAspHisLysThrThrProCysSerSerAl 417
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1425 TCTCAAAGGGCTGCATTCTGTGATCACAAGACCACTCCATGCTCTTCAGC 1474

 417 aIleIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeuGlnProG 434
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1475 AATAATAAATCCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAGCCTG 1524

 434 lyIleAlaGlnGlnTrpIleGlnSerLysArgGluAspIleValAsnGln 450
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1525 GTATAGCCCAGCAGTGGATCCAGAGCAAAAGGGAAGACATTGTGAACCAA 1574

 451 MetThrGluAlaCysLeuAsnGlnSerLeuAspAlaLeuLeuSerArgAs 467
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1575 ATGACAGAAGCCTGCCTTAACCAGTCGCTAGATGCCCTTCTGTCCAGGGA 1624

 467 pLeuIleMetLysGluAspTyrGluLeuValSerThrLysProThrArgT 484
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1625 CTTGATCATGAAAGAGGACTATGAACTTGTTAGTACCAAGCCTACAAGGA 1674

 484 hrSerLysValArgGlnLeuLeuAspThrThrAspIleGlnGlyGluGlu 500
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1675 CCTCAAAAGTCAGACAATTACTAGACACTACTGACATCCAAGGAGAAGAA 1724

 501 PheAlaLysValIleValGlnLysLeuLysAspAsnLysGlnMetGlyLe 517
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1725 TTTGCCAAAGTTATAGTACAAAAATTGAAAGATAACAAACAAATGGGTCT 1774

 517 uGlnProTyrProGluIleLeuValValSerArgSerProSerLeuAsnL 534
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1775 TCAGCCTTACCCGGAAATACTTGTGGTTTCTAGATCACCATCTTTAAATT 1824

 534 euLeuGlnAsnLysSerMet 540
     ||||||||||||||||||||
1825 TACTTCAAAATAAAAGCATG 1844

seq_name: gb_pr1:AC004003

seq_documentation_block:
LOCUS       AC004003   116650 bp    DNA             PRI       21-DEC-1999
DEFINITION  Homo sapiens BAC clone CTA-437L15 from 8q21, complete sequence.
ACCESSION   AC004003
VERSION     AC004003.1  GI:2772557
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 116650)
  AUTHORS   Ozersky,P., Holmes,A. and Broy,M.
  TITLE     The sequence of Homo sapiens BAC clone CTA-437L15
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 116650)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-1998) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   3  (bases 1 to 116650)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JUN-1998) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   4  (bases 1 to 116650)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
```

```
JOURNAL     Submitted (21-DEC-1999) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_RG437L15
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded or sequenced with an alternate
            chemistry; an attempt was made to resolve all sequencing problems,
            such as compressions and repeats; all regions were covered by
            sequence from more than one subclone; and the assembly was
            confirmed by restriction digest.

            MAPPING INFORMATION:
            This chromosome 8 clone was provided by Dr. Patrick Concannon
            (patcon@vmmc.org) at the Virginia Mason Research Institute.

            SOURCE INFORMATION:
            Clone CTA-437L15 is from a release of the human BAC library
            CITB-HS-A.  The library contains cloned DNA from human sperm.  See:
             Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J.
            Kim et al., Genomics 34:213-8 (1996).  The clone is available from
            Research Genetics, Inc. (http://www.resgen.com).
            VECTOR:  pBeloBAC11
            Selection:  chloramphenicol

            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is CTA-237G1. The actual start of
            this clone is at base position 1 of CTA-437L15; actual end is at
            116650 of CTA-437L15.

            This clone contains STS HS275YF1 (NID:g1051703).
FEATURES             Location/Qualifiers
     source          1. .116650
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8q21"
                     /clone="CTA-437L15"
                     /clone_lib="CITB-HS-A"
     repeat_region   150. .530
                     /rpt_family="L1"
     repeat_region   498. .649
                     /rpt_family="L1"
     repeat_region   673. .1096
                     /rpt_family="MaLR"
     repeat_region   3188. .3348
                     /rpt_family="L1"
     repeat_region   3605. .3726
                     /rpt_family="(TA)n"
     repeat_region   3845. .4563
                     /rpt_family="L1"
     repeat_region   5641. .6075
                     /rpt_family="Retroviral"
     repeat_region   6133. .6433
                     /rpt_family="Alu"
     misc_feature    7110. .8040
                     /note="CpG_island (%GC=69.1, o/e=0.80, #CpGs=93)"
     gene            7872. .40218
                     /gene="WUGSC:H_RG437L15.1"
     CDS             join(7872. .8044,12642. .12795,15154. .15309,19584. .19741,
                     21552. .21601,22478. .22639,29890. .29975,33855. .33944,
                     36395. .36488,39123. .39284,39881. .40218)
                     /gene="WUGSC:H_RG437L15.1"
                     /note="serine/threonine kinase RICK; match to protein
                     AF027706 (PID:g3123887) and mRNA AF027706 (NID:g3123886);
                     H_RG437L15.1"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="AAC24561.1"
                     /db_xref="GI:3264574"
                     /translation="MNGEAICSALPTIPYHKLADLRYLSRGASGTVSSARHADWRVQV
                     AVKHLHIHTPLLDSERKDVLREAEILHKARFSYILPILGICNEPEFLGIVTEYMPNGS
                     LNELLHRKTEYPDVAWPLRFRILHEIALGVNYLHNMTPPLLHHDLKTQNILLDNEFHV
                     KIADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPENYEPGQKSRASIKHDIYSYAVI
                     TWEVLSRKQPFEDVTNPLQIMYSVSQGHRPVINEESLPYDIPHRARMISLIESGWAQN
                     PDERPSFLKCLIELEPVLRTFEEITFLEAVIQLKKTKLQSVSSAIHLCDKKKMELSLN
                     IPVNHGPQEESCGSSQLHENSGSPETSRSLPAPQDNDFLSRKAQDCYFMKLHHCPGNH
                     SWDSTISGSQRAAFCDHKTTPCSSAIINPLSTAGNSERLQPGIAQQWIQSKREDIVNQ
                     MTEACLNQSLDALLSRDLIMKEDYELVSTKPTRTSKVRQLLDTTDIQGEEFAKVIVQK
                     LKDNKQMGLQPYPEILVVSRSPSLNLLQNKSM"
     repeat_region   8463. .8770
                     /rpt_family="MER1_type"
     repeat_region   10435. .10540
                     /rpt_family="MIR"
     repeat_region   11433. .11538
                     /rpt_family="Alu"
     repeat_region   11790. .12003
                     /rpt_family="Alu"
     repeat_region   12466. .12557
                     /rpt_family="MIR"
     repeat_region   15579. .16292
                     /rpt_family="L1"
     repeat_region   16292. .16429
                     /rpt_family="L1"
     repeat_region   16460. .16991
                     /rpt_family="L1"
     repeat_region   22863. .23483
                     /rpt_family="L2"
     repeat_region   23575. .23750
                     /rpt_family="L1"
     repeat_region   23761. .23924
                     /rpt_family="Alu"
     repeat_region   23929. .24439
                     /rpt_family="L1"
     repeat_region   24471. .24960
                     /rpt_family="L2"
     repeat_region   24967. .25285
                     /rpt_family="Alu"
     repeat_region   25337. .25671
                     /rpt_family="L2"
     repeat_region   26151. .26280
                     /rpt_family="MER1_type"
     repeat_region   26268. .26500
                     /rpt_family="MER1_type"
     repeat_region   26611. .27000
                     /rpt_family="L1"
     repeat_region   28128. .29109
                     /rpt_family="L1"
     repeat_region   29128. .29434
                     /rpt_family="Alu"
     repeat_region   29435. .29752
                     /rpt_family="L1"
     repeat_region   30205. .30341
                     /rpt_family="Alu"
     repeat_region   31170. .31344
                     /rpt_family="Alu"
     repeat_region   31466. .31688
                     /rpt_family="MER2_type"
     repeat_region   31761. .31955
                     /rpt_family="Alu"
     repeat_region   31956. .32225
                     /rpt_family="MER2_type"
     repeat_region   32329. .32460
                     /rpt_family="Alu"
     repeat_region   33155. .33281
```

```
                    /rpt_family="MIR"
     repeat_region  33380..33437
                    /rpt_family="MIR"
     repeat_region  34433..34514
                    /rpt_family="MER2_type"
     repeat_region  34525..34939
                    /rpt_family="MaLR"
     repeat_region  34999..35256
                    /rpt_family="Alu"
     repeat_region  35388..35477
                    /rpt_family="MER2_type"
     repeat_region  35994..36286
                    /rpt_family="Alu"
     repeat_region  36686..36985
                    /rpt_family="Alu"
     repeat_region  37488..37673
                    /rpt_family="L2"
     repeat_region  37862..38161
                    /rpt_family="Alu"
     repeat_region  38297..38599
                    /rpt_family="Alu"
     repeat_region  41954..42006
                    /rpt_family="L2"
     repeat_region  42596..42777
                    /rpt_family="MER1_type"
     repeat_region  43357..43658
                    /rpt_family="Alu"
     repeat_region  44088..44386

alignment_scores:
           Quality:  726.00             Length:    364
             Ratio:  4.400                Gaps:      2
 Percent Similarity:  45.330   Percent Identity: 45.055

alignment_block:
 US-09-445-223-1 x AC004003   ..

 Align seg 1/1  to: AC004003  from: 1  to: 116650

     376 LysAlaGlnAspCysTyrPheMetLysLeuHisHisCysProGlyAsnHi 392
         ||||||||||||||||||||||||||||||||||||||||||||||||||
   39125 AAAGCTCAAGACTGTTATTTTATGAAGCTGCATCACTGTCCTGGAAATCA 39174

     392 sSerTrpAspSerThrIleSerGlySerGlnArgAlaAlaPheCysAspH 409
         ||||||||||||||||||||||||||||||||||||||||||||||||||
   39175 CAGTTGGGATAGCACCATTTCTGGATCTCAAAGGGCTGCATTCTGTGATC 39224

     409 isLysThrThrProCysSerSerAlaIleIleAsnProLeuSerThrAla 425
         ||||||||||||||||||||||||||||||||||||||||||||||||||
   39225 ACAAGACCACTCCATGCTCTTCAGCAATAATAAATCCACTCTCAACTGCA 39274

     426 GlyAsnSer......................................... 428
         |||||||||
   39275 GGAAACTCAGGTAAATATTTACTGTGAAACACCTTGTAAGTGATGCCATT 39324

     428 .................................................. 428

   39325 GACCTTACATATCTGTGTTCAGTGGATTTACCATACAATGAGGTTTGTCT 39374

     428 .................................................. 428

   39375 TTATTTCCTCAATATGATAAATACATACCTTCTTAAGTAGGAAAACAAAT 39424

     428 .................................................. 428

   39425 CCCAGGAGATGGGAATGATGACATGGCCCAAATGGTATTCTATGAACATC 39474

     428 .................................................. 428

   39475 TAAATAGGGAGGCAGTGCTAAGAGGAGGTATTAACATATTACTAATGGTA 39524

     428 .................................................. 428

   39525 AAGTTATTTTTGTTTATTTAACAAATTTGCAGGTATCTTTTATATTTGAA 39574

     428 .................................................. 428

   39575 ATAAAATGGACATCAAACAAAACTGAACAAATCCCAGAAGATGGGAATGA 39624

     428 .................................................. 428

   39625 TGACATAAAGTGGTTTTTGCCCCGTGATTTATACCATAGGAAAAGGCAGT 39674

     428 .................................................. 428

   39675 GCCTTTTCTGGGGAGTGAGGGGGAGATAACCCAAATCTCTGAATGTTTCC 39724

     428 .................................................. 428

   39725 TGAAAAATGACACCATTCCCCAAAGTGTTGGTGTTCCATATTGGCACAAT 39774

     428 .................................................. 428

   39775 CATAGGAAAATTATGTTTTATGCTTTATTTATTTTTCATGGAATTTTACT 39824

     428 .................................................. 428

   39825 TATTGCTATGTATCTGTCCTCACCTTTTAAATTATTCATTCCTTGTTCTT 39874

     429 ...Glu..ArgLeuGlnProGlyIleAlaGlnGlnTrpIleGlnSerLys 443
            :::  ||||||||||||||||||||||||||||||||||||||||||
   39875 TTTCAGAACGTCTGCAGCCTGGTATAGCCCAGCAGTGGATCCAGAGCAAA 39924

     444 ArgGluAspIleValAsnGlnMetThrGluAlaCysLeuAsnGlnSerLe 460
         ||||||||||||||||||||||||||||||||||||||||||||||||||
   39925 AGGGAAGACATTGTGAACCAAATGACAGAAGCCTGCCTTAACCAGTCGCT 39974

     460 uAspAlaLeuLeuSerArgAspLeuIleMetLysGluAspTyrGluLeuV 477
         ||||||||||||||||||||||||||||||||||||||||||||||||||
   39975 AGATGCCCTTCTGTCCAGGGACTTGATCATGAAAGAGGACTATGAACTTG 40024

     477 alSerThrLysProThrArgThrSerLysValArgGlnLeuLeuAspThr 493
         ||||||||||||||||||||||||||||||||||||||||||||||||||
   40025 TTAGTACCAAGCCTACAAGGACCTCAAAAGTCAGACAATTACTAGACACT 40074

     494 ThrAspIleGlnGlyGluGluPheAlaLysValIleValGlnLysLeuLy 510
         ||||||||||||||||||||||||||||||||||||||||||||||||||
   40075 ACTGACATCCAAGGAGAAGAATTTGCCAAAGTTATAGTACAAAAATTGAA 40124

     510 sAspAsnLysGlnMetGlyLeuGlnProTyrProGluIleLeuValValS 527
         ||||||||||||||||||||||||||||||||||||||||||||||||||
   40125 AGATAACAAACAAATGGGTCTTCAGCCTTACCCGGAAATACTTGTGGTTT 40174

     527 erArgSerProSerLeuAsnLeuLeuGlnAsnLysSerMet 540
         |||||||||||||||||||||||||||||||||||||||||
   40175 CTAGATCACCATCTTTAAATTTACTTCAAAATAAAGCATG 40215

seq_name: gb_pr4:AF117829

seq_documentation_block:
LOCUS       AF117829   320250 bp    DNA             PRI       13-JAN-1999
DEFINITION  Homo sapiens 8q21.3: RICK gene, complete sequence.
ACCESSION   AF117829
VERSION     AF117829.1  GI:4151947
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 320250)
  AUTHORS   Platzer,M. and Varon,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-DEC-1998) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
COMMENT     This sequence is part of a larger genomic contig. The start of this
            sequence is directed towards the centromere. The end of this
```

```
            sequence (320001. .320250) overlaps with the start of the
            neighbouring Acc_number AF049895 (1. .250).
FEATURES             Location/Qualifiers
     source          1. .320250
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8q21.3"
     source          1. .137184
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="P64M4"
     repeat_region   complement(572. .704)
                     /rpt_family="MIR"
     repeat_region   705. .870
                     /rpt_family="MIR"
     repeat_region   complement(930. .985)
                     /rpt_family="MIR"
     repeat_region   2222. .2456
                     /rpt_family="L2"
     repeat_region   2538. .2661
                     /rpt_family="L2"
     repeat_region   complement(2695. .2845)
                     /rpt_family="FAM"
     repeat_region   2913. .3485
                     /rpt_family="L2"
     repeat_region   3559. .3679
                     /rpt_family="L2"
     repeat_region   3776. .3930
                     /rpt_family="MIR"
     repeat_region   complement(4035. .4247)
                     /rpt_family="MIR"
     repeat_region   complement(5211. .5655)
                     /rpt_family="L2"
     repeat_region   5997. .6100
                     /rpt_family="L2"
     repeat_region   6800. .6850
                     /rpt_family="(TAGA)n"
     repeat_region   8389. .8709
                     /rpt_family="L2"
     repeat_region   complement(9413. .9761)
                     /rpt_family="MLT1A1"
     repeat_region   10399. .10473
                     /rpt_family="L2"
     repeat_region   11341. .11507
                     /rpt_family="MIR"
     repeat_region   complement(11525. .11570)
                     /rpt_family="(TAA)n"
     repeat_region   11931. .11985
                     /rpt_family="(CATG)n"
     repeat_region   complement(12051. .12149)
                     /rpt_family="MER53"
     repeat_region   complement(12351. .12760)
                     /rpt_family="MSTA"
     repeat_region   14544. .14734
                     /rpt_family="MIR"
     repeat_region   complement(15740. .15847)
                     /rpt_family="MIR"
     repeat_region   16491. .16558
                     /rpt_family="MIR"
     repeat_region   18103. .18162
                     /rpt_family="MIR"
     repeat_region   18343. .18639
                     /rpt_family="AluSx"
     repeat_region   19005. .19075
                     /rpt_family="MER57-internal"
     repeat_region   complement(19478. .19560)
                     /rpt_family="MIR"
     repeat_region   complement(23416. .23710)
                     /rpt_family="AluSg"
     repeat_region   complement(24833. .25025)
                     /rpt_family="MIR"
     repeat_region   complement(25333. .25392)
                     /rpt_family="MER81"
     repeat_region   25969. .26053
                     /rpt_family="L1MB8"
     repeat_region   26167. .26302
                     /rpt_family="L1M4"
     repeat_region   complement(26512. .26684)
                     /rpt_family="AluJo"
     repeat_region   26685. .26712
                     /rpt_family="(TA)n"
     repeat_region   complement(26713. .26841)
                     /rpt_family="AluJo"
     repeat_region   complement(27456. .27620)
                     /rpt_family="MER5A"
     repeat_region   complement(27724. .28020)
                     /rpt_family="AluY"
     repeat_region   28147. .28283
                     /rpt_family="AluY"
     repeat_region   28284. .28311
                     /rpt_family="(TA)n"
     repeat_region   28312. .28476
                     /rpt_family="AluY"
     repeat_region   29222. .29518
                     /rpt_family="L1M2"
     repeat_region   29454. .32583
                     /rpt_family="L1MA6"
     repeat_region   32636. .33403
                     /rpt_family="L1MA6"
     repeat_region   33404. .33705
                     /rpt_family="AluSc"
     repeat_region   33706. .33960
                     /rpt_family="L1MA6"
     repeat_region   complement(34346. .34614)
                     /rpt_family="MER8"
     repeat_region   34786. .34873
                     /rpt_family="MIR"
     repeat_region   34839. .35057
                     /rpt_family="L2"
     repeat_region   complement(35173. .35382)
                     /rpt_family="AluYa5"
     repeat_region   35577. .35746
                     /rpt_family="MLT1D"
     repeat_region   complement(35747. .36046)
                     /rpt_family="AluY"
     repeat_region   36047. .36323
                     /rpt_family="MLT1D"
     repeat_region   complement(36593. .36761)
                     /rpt_family="MIR"
     repeat_region   complement(37858. .37899)
                     /rpt_family="L2"
     repeat_region   38652. .38895
                     /rpt_family="L1M2"
     repeat_region   38959. .39811
                     /rpt_family="L1M3e"
     repeat_region   39885. .39989
                     /rpt_family="L1M2"
     repeat_region   40242. .43133
                     /rpt_family="L1M2"
     repeat_region   43134. .43499
                     /rpt_family="MLT1A1"
     repeat_region   43500. .44075
                     /rpt_family="L1M2"
     repeat_region   44076. .44391
                     /rpt_family="AluJo"
     repeat_region   44392. .45758
                     /rpt_family="L1MA4A"
     repeat_region   complement(45759. .46216)
                     /rpt_family="L1PA7"
     repeat_region   46217. .47676
                     /rpt_family="L1MA4A"
     repeat_region   47693. .47955
                     /rpt_family="L1MA5A"
     repeat_region   48227. .48421
                     /rpt_family="L2"
```

```
  repeat_region   complement(49170. .49277)
                  /rpt_family="MIR"
  repeat_region   complement(49409. .50232)
                  /rpt_family="L1PA2"
  repeat_region   50233. .50476
                  /rpt_family="L1PA4"
  repeat_region   complement(50588. .50644)
                  /rpt_family="MIR"
  repeat_region   complement(51729. .52082)
                  /rpt_family="THE1C"
  repeat_region   complement(52083. .53629)
                  /rpt_family="THE1C-internal"
  repeat_region   complement(53632. .53992)
                  /rpt_family="THE1C"
  repeat_region   54101. .54351
                  /rpt_family="L2"
  repeat_region   54433. .54822
                  /rpt_family="L2"
  repeat_region   54823. .55177
                  /rpt_family="MLT1A2"
  repeat_region   55178. .55291
                  /rpt_family="L2"
  repeat_region   complement(55316. .55371)
                  /rpt_family="L2"
  repeat_region   55545. .55878
                  /rpt_family="L2"
  repeat_region   57436. .57481
                  /rpt_family="(TTAAA)n"
  repeat_region   60524. .60721
                  /rpt_family="L1ME"
  repeat_region   complement(61481. .61865)
                  /rpt_family="MSTA"
  repeat_region   62004. .62295
                  /rpt_family="MLT1E"
  repeat_region   62427. .67614
                  /rpt_family="L1PA2"
  source          62833. .145579
                  /organism="Homo sapiens"

alignment_scores:
          Quality:  726.00          Length:    364
            Ratio:  4.400             Gaps:      2
 Percent Similarity:  45.330   Percent Identity: 45.055

alignment_block:
 US-09-445-223-1 x AF117829   ..

 Align seg 1/1  to: AF117829  from: 1  to: 320250

     376 LysAlaGlnAspCysTyrPheMetLysLeuHisHisCysProGlyAsnHi 392
         ||||||||||||||||||||||||||||||||||||||||||||||||||
  178694 AAAGCTCAAGACTGTTATTTTATGAAGCTGCATCACTGTCCTGGAAATCA 178743

     392 sSerTrpAspSerThrIleSerGlySerGlnArgAlaAlaPheCysAspH 409
         ||||||||||||||||||||||||||||||||||||||||||||||||||
  178744 CAGTTGGGATAGCACCATTTCTGGATCTCAAAGGGCTGCATTCTGTGATC 178793

     409 isLysThrThrProCysSerSerAlaIleIleAsnProLeuSerThrAla 425
         ||||||||||||||||||||||||||||||||||||||||||||||||||
  178794 ACAAGACCACTCCATGCTCTTCAGCAATAATAAATCCACTCTCAACTGCA 178843

     426 GlyAsnSer......................................... 428
         |||||||||
  178844 GGAAACTCAGGTAAATATTTACTGTGAAACACCTTGTAAGTGATGCCATT 178893

     428 .................................................. 428

  178894 GACCTTACATATCTGTGTTCAGTGGATTTACCATACAATGAGGTTTGTCT 178943

     428 .................................................. 428

  178944 TTATTTCCTCAATATGATAAATACATACCTTCTTAAGTAGGAAAACAAAT 178993
```

```
     428 .................................................. 428

  178994 CCCAGGAGATGGGAATGATGACATGGCCCAAATGGTATTCTATGAACATC 179043

     428 .................................................. 428

  179044 TAAATAGGGAGGCAGTGCTAAGAGGAGGTATTAACATATTACTAATGGTA 179093

     428 .................................................. 428

  179094 AAGTTATTTTTGTTTATTTAAAAAATTTGCAGGTATCTTTTATATTTGAA 179143

     428 .................................................. 428

  179144 ATAAAATGGACATCAAACAAAACTGAACAAATCCCAGAAGATGGAAATGA 179193

     428 .................................................. 428

  179194 TGACATAAAGTGGTTTTTGCCCCGTGATTTATACCATAGGAAAAGGCAAT 179243

     428 .................................................. 428

  179244 GCCTTTTCTGGGGAGTGAGGGGGAGATAACCCAAATCTCTGAATGTTTCC 179293

     428 .................................................. 428

  179294 TGAAAAATGACACCATTCCCCAAAGTGTTGGTGTTCCATATTGGCACAAT 179343

     428 .................................................. 428

  179344 CATAGGAAAATTATGTTTTATGCTTTATTTATTTTTCATGGAATTTTACT 179393

     428 .................................................. 428

  179394 TATTGCTATGTATCTGTCCTCACCTTTTAAATTATTCATTCCTTGTTCTT 179443

     429 ...Glu..ArgLeuGlnProGlyIleAlaGlnGlnTrpIleGlnSerLys 443
            :::  ||||||||||||||||||||||||||||||||||||||||||
  179444 TTTCAGAACGTCTGCAGCCTGGTATAGCCCAGCAGTGGATCCAGAGCAAA 179493

     444 ArgGluAspIleValAsnGlnMetThrGluAlaCysLeuAsnGlnSerLe 460
         ||||||||||||||||||||||||||||||||||||||||||||||||||
  179494 AGGGAAGACATTGTGAACCAAATGACAGAAGCCTGCCTTAACCAGTCGCT 179543

     460 uAspAlaLeuLeuSerArgAspLeuIleMetLysGluAspTyrGluLeuV 477
         ||||||||||||||||||||||||||||||||||||||||||||||||||
  179544 AGATGCCCTTCTGTCCAGGGACTTGATCATGAAAGAGGACTATGAACTTG 179593

     477 alSerThrLysProThrArgThrSerLysValArgGlnLeuLeuAspThr 493
         ||||||||||||||||||||||||||||||||||||||||||||||||||
  179594 TTAGTACCAAGCCTACAAGGACCTCAAAAGTCAGACAATTACTAGACACT 179643

     494 ThrAspIleGlnGlyGluGluPheAlaLysValIleValGlnLysLeuLy 510
         ||||||||||||||||||||||||||||||||||||||||||||||||||
  179644 ACTGACATCCAAGGAGAAGAATTTGCCAAAGTTATAGTACAAAAATTGAA 179693

     510 sAspAsnLysGlnMetGlyLeuGlnProTyrProGluIleLeuValValS 527
         ||||||||||||||||||||||||||||||||||||||||||||||||||
  179694 AGATAACAAACAAATGGGTCTTCAGCCTTACCCGGAAATACTTGTGGTTT 179743

     527 erArgSerProSerLeuAsnLeuLeuGlnAsnLysSerMet 540
         |||||||||||||||||||||||||||||||||||||||||
  179744 CTAGATCACCATCTTTAAATTTACTTCAAAATAAAAGCATG 179784

seq_name: gb_rol:AF302127

seq_documentation_block:
LOCUS       AF302127     3559 bp    mRNA            ROD       08-NOV-2000
DEFINITION  Mus musculus PKC-regulated kinase PKK mRNA, complete cds.
ACCESSION   AF302127
VERSION     AF302127.1  GI:11120430
KEYWORDS    .
SOURCE      house mouse.
```

```
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3559)
  AUTHORS   Chen,L., Haider,K., Cariappa,A., Rowitch,D. and Pillai,S.
  TITLE     PKK, a novel ankyrin repeat-containing protein kinase regulated by
            protein kinase C-beta
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3559)
  AUTHORS   Chen,L., Haider,K., Cariappa,A. and Pillai,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-SEP-2000) Cancer Center, Massachusetts General
            Hospital, 149 13th Street, Charlestown, MA 02129-2060, USA
FEATURES             Location/Qualifiers
     source          1. .3559
                     /organism="Mus musculus"
                     /strain="BALB/c"
                     /db_xref="taxon:10090"
     CDS             50. .2410
                     /codon_start=1
                     /product="PKC-regulated kinase PKK"
                     /protein_id="AAG30871.1"
                     /db_xref="GI:11120431"
                     /translation="MEGEGRGRWALGLLRTFDAGEFAGWEKVGSGGFGQVYKVRHVHW
                     KTWLAIKCSPSLHVDDRERMELLEEAKKMEMAKFRYILPVYRICQEPVGLVMEYMETG
                     SLEKLLASEPLPWDLRFRIVHETAVGMNFLHCMSPPLLHLDLKPANILLDAHYHVKIS
                     DFGLAKCNGMSHSQDLSMDGLFGTIAYLPPERIREKSRLFDTKHDVYSFAIVIWGVLT
                     QKKPFADEKNILHIMMKVVKGHRPELPPICRPRPRACASLIGLMQRCWHADPQVRPTF
                     QEITSETEDLCEKPDEEVKDLAHEPGEKSSLESKSEARPESSRLKRASAPPFDNDCSL
                     SELLSQLDSGISQTLEGPEELSRSSSECKLPSSSSGKRLSGVSSVDSAFSSRGSLSLS
                     FEREASTGDLGPTDIQKKKLVDAIISGDTSRLMKILQPQDVDLVLDSSASLLHLAVEA
                     GQEECVKWLLLNNANPNLTNRKGSTPLHMAVERKGRGIVELLLARKTSVNAKDEDQWT
                     ALHFAAQNGDEASTRLLLEKNASVNEVDFEGRTPMHVACQHGQENIVRTLLRRGVDVG
                     LQGKDAWLPLHYAAWQGHLPIVKLLAKQPGVSVNAQTLDGRTPLHLAAQRGHYRVARI
                     LIDLCSDVNICSLQAQTPLHVAAETGHTSTARLLLHRGAGKEALTSEGYTALHLAAQN
                     GHLATVKLLIEEKADVMARGPLNQTALHLAAARGHSEVVEELVSADLIDLSDEQGLSA
                     LHLAAQGRHSQTVETLLKHGAHINLQSLKFQGGQSSAATLLRRSKT"
BASE COUNT      826 a    901 c   1000 g    832 t
ORIGIN


alignment_scores:
           Quality:  577.00             Length:    553
             Ratio:  1.930                Gaps:     27
 Percent Similarity:  54.069   Percent Identity:  33.816

alignment_block:
US-09-445-223-1 x AF302127   ..

 Align seg 1/1  to: AF302127  from: 1  to: 3559

   1 MetAsnGlyGluAla............IleCysSerAlaLeuProThrIl 13
     |||:::||||||:::                  :::   |||   |||::
  50 ATGGAGGGCGAGGGCCGGGGCCGGTGGGCTCTGGGGCTGCTGCGCACCTT 99

  13 eProTyrHisLysLeuAlaAspLeuArgTyrLeuSerArgGlyAlaSerG 30
     :         ::::::|||      :::   ::::::   |||:::   |
 100 CGACGCCGGCGAATTCGCAGGCTGGGAGAAGGTGGGCTCGGGCGGCTTCG 149

  30 lyThrValSerSerAlaArgHisAlaAspTrpArgValGlnValAlaVal 46
     ||   |||   ::::::||||||:::   |||::::::   :::|||:::
 150 GGCAGGTGTACAAGGTGCGCCATGTGCACTGGAAGACGTGGCTCGCGATC 199

  47 Lys.........HisLeuHisIleHisThrProLeuLeuAspSerGluAr 60
     |||            ||||||:::               |||   |||||
 200 AAGTGCTCGCCCAGTCTGCACGTCGAC............GACAGGGAACG 237

  60 gLysAspValLeuArgGluAlaGluIleLeuHisLysAlaArgPheSerT 77
     |   ::::::|||:::||||||:::   ::::::   |||:::|||   |
 238 AATGGAGCTCCTGGAGGAAGCTAAGAAGATGGAGATGGCCAAGTTCCGAT 287

  77 yrIlePheProIleLeuGlyIleCysAsnGluProGluPheLeuGlyIle 93
     |||||:::|||:::      ||||||:::||||||      :::|||:::
```

```
 288 ACATTCTACCTGTGTACCGCATATGCCAGGAACCT......GTCGGCTTG 331

  94 ValThrGluTyrMetProAsnGlySerLeuAsnGluLeuLeuHisArgLy 110
     |||   |||||||||   :::|||||||||::::::||||||      ::
 332 GTCATGGAGTACATGGAGACAGGCTCCCTGGAAAAGCTGCTGGCCTCAGA 381

 110 sThrGluTyrProAspValAlaTrpProLeuArgPheArgIleLeuHisG 127
     :               :::   |||   |||||||||||||||:::||||
 382 G.............CCATTGCCTTGGACCTGCGCTTTCGCATCGTGCACG 419

 127 luIleAlaLeuGlyValAsnTyrLeuHisAsnMetThrProProLeuLeu 143
     ||   |||:::|||:::|||:::||||||   |||:::||||||||||||
 420 AGACAGCCGTGGGCATGAACTTCCTGCATTGCATGTCTCCGCCACTGCTG 469

 144 HisHisAspLeuLysThrGlnAsnIleLeuLeuAspAsnGluPheHisVa 160
     |||   |||||||||      |||||||||||||||   ::::::|||||
 470 CACCTAGACCTGAAGCCAGCGAACATCCTGCTGGATGCCCACTACCATGT 519

 160 lLysIleAlaAspPheGlyLeuSerLysTrpArgMetMetSerLeuSerG 177
     |||||||:::||||||||||||:::|||   :::   ||||||   ||||
 520 CAAGATTTCTGACTTTGGGCTGGCCAAGTGCAATGGCATGTCCCACTCTC 569

 177 lnSerArgSerSerLysSerAlaProGluGly......GlyThrIleIle 191
     ||:::   |||               :::|||      |||||||||
 570 AGGACCTCAGCATG............GATGGCCTGTTTGGTACAATCGCT 607

 192 TyrMetProProGluAsnTyrGluProGlyGlnLysSerArgAla...Se 207
     |||:::|||||||||:::   :::      :::|||||||||      ::
 608 TACCTCCCTCCAGAGCGAATTCGT......GAGAAGAGCCGCTTGTTTGA 651

 207 rIleLysHisAspIleTyrSerTyrAlaValIleThrTrpGluValLeuS 224
     :   |||||||||:::||||||:::|||::::::   |||   ||||||:
 652 CACCAAACATGATGTATACAGCTTCGCCATTGTGATCTGGGGTGTGCTTA 701

 224 erArgLysGlnProPheGluAspValThrAsnProLeuGlnIleMetTyr 240
     :::::|||:::||||||   |||      |||   |||:::||||||
 702 CACAGAAGAAGCCATTTGCAGATGAAAAGAACATCCTACACATCATGATG 751

 241 SerValSerGlnGlyHisArgProValIleAsnGluGluSerLeuProTy 257
     :::|||   :::||||||||||||            |||      |||
 752 AAAGTGGTAAAGGGCCACCGCCCA............GAGCTGCCACCCAT 789

 257 rAspIleProHisArg.........AlaArgMetIleSerLeuIleGluS 271
            |||:::            |||   :::|||:::|||::::::
 790 CTGCAGACCCCGGCCGCGTGCCTGTGCCAGCCTGATAGGGCTCATGCAAC 839

 271 erGlyTrpAlaGlnAsnProAspGluArgProSerPheLeuLysCysLeu 287
          |||      :::|||:::   ||||||
 840 GGTGCTGGCATGCAGACCCACAGGTGCGGCCC.................. 871

 288 IleGluLeuGluProValLeuArgThrPheGluGluIleThrPheLeuGl 304
                             ||||||:::|||||||||
 872 ........................ACCTTCCAAGAAATTACC........ 889

 304 uAlaValIleGlnLeuLysLysThrLysLeuGlnSerValSerSerAlaI 321
                                             |||::::::
 890 ........................................TCTGAAACAG 899

 321 leHisLeuCysAspLysLysLysMetGluLeuSerLeuAsnIleProVal 337
          ||||||:::|||         |||::::::            :::
 900 AAGACCTTTGTGAGAAGCCTGATGAGGAGGTGAAA.........GACCTG 940

 338 AsnHisGlyProGlnGluGluSerCysGlySerSerGlnLeuHisGluAs 354
        |||   |||   |||:::|||      :::|||:::   :::   ::
 941 GCTCATGAGCCAGGCGAGAAAAGCTCTCTAGAGTCCAAGAGTGAGGCCAG 990

 354 nSerGlySerProGluThrSerArgSerLeuProAlaProGlnAspAsnA 371
     :      |||   :::   :::|||:::         |||   |||||||
 991 GCCCGAGTCCTCACGCCTCAAGCGCGCCTCTGCTCCCCCCTTCGATAACG 1040

 371 sp.PheLeuSerArgLysAlaGlnAspCysTyrPheMetLysLeuHisHi 387
     ||    :::|||            :::|||
1041 ACTGCAGTCTCT.........CCGAGTTGC.................... 1061
```

```
 387 sCysProGlyAsnHisSerTrpAspSerThrIleSerGlySerGlnArgA 404
      |||         |||||||||                ||||||   |||
1062 .TGT.........CACAGTTGG..........ACTCTGGGATCTCCCAGAC 1092

 404 laAlaPheCysAspHisLysThrThrProCysSerSerAlaIleIleAsn 420
                             :::|||   |||||||||   :::
1093 TCTTG...............AAGGCCCCGAAGAGCTCAGCCGAAGTT... 1124

 421 ProLeuSerThrAlaGlyAsnSerGluArgLeuGlnProGlyIleAlaGl 437
     ||||||            |||::::::      :::|||::::::|||::
1125 CCTCTG............AATGCAAGCTCCCATCGTCCAGCAGTGGCAAG 1162

 437 n..................GlnTrpIleGlnSerLysArgGluAspIleV 448
     :                  ||||||   |||             :::
1163 AGGCTCTCGGGGGTGTCCTCAGTGGACTCAGCCTTTTCCTCCAGAG.... 1208

 448 alAsnGlnMetThrGluAlaCysLeuAsnGlnSerLeuAspAlaLeuLeu 464
                   :::   |||                 :::|||||||
1209 ..............GATCGCTGT...............CACTGTCTTTTG 1229

 465 SerArgAspLeuIleMetLysGluAspTyrGluLeuValSerThrLysPr 481
     |||
1230 AGC............................................... 1232

 481 oThrArgThrSerLysValArgGlnLeu..LeuAspThrThrAspIleGl 497
               :::|||::::::|||     |||      |||||||||||
1233 ..........GGGAAGCTTCAACAGGCGACCTGGGCCCCACAGACATCCA 1272

 497 nGlyGluGluPheAlaLysValIleValGlnLysLeuLysAspAsnLysG 514
     |   ::::::::::::   :::|||::::::         ::::::   :
1273 GAAGAAGAAGCTAGTGGATGCCATCATATCAGGGGACACCAGCAGGCTGA 1322

 514 lnMetGlyLeuGlnProTyrProGluIleLeuValValSerArgSerPro 530
     ::      |||||||||            ||||||::::::   |||
1323 TGAAGATCCTACAGCCCCAAGATGTGGACTTGGTTCTAGACAGCAGTGCC 1372

 531 SerLeu 532
     ||||||
1373 AGCCTG 1378

seq_name: gb_pr8:HSA278016

seq_documentation_block:
LOCUS       HSA278016     3879 bp    mRNA            PRI       19-NOV-2000
DEFINITION  Homo sapiens mRNA for protein kinase (dik gene).
ACCESSION   AJ278016
VERSION     AJ278016.1  GI:9886710
KEYWORDS    dik gene; protein kinase.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3879)
  AUTHORS   Bahr,C., Rohwer,A., Stempka,L., Rincke,G., Marks,F. and
            Gschwendt,M.
  TITLE     DIK, a Novel Protein Kinase That Interacts with Protein Kinase
            Cdelta. CLONING, CHARACTERIZATION, AND GENE ANALYSIS
  JOURNAL   J. Biol. Chem. 275 (46), 36350-36357 (2000)
   PUBMED   10948194
REFERENCE   2  (bases 1 to 3879)
  AUTHORS   Gschwendt,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-MAY-2000) Gschwendt M., Biochemistry of
            Tissue-specific Regulation, German Cancer Research Center, Im
            Neuenheimer Feld 280, D-69151 Heidelberg, GERMANY
FEATURES             Location/Qualifiers
     source          1. .3879
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HaCaT"
     gene            49. .3843
                     /gene="dik"
```

```
     CDS             49. .2403
                     /gene="dik"
                     /function="protein phosphorylation"
                     /codon_start=1
                     /evidence=experimental
                     /product="protein kinase"
                     /protein_id="CAC04247.1"
                     /db_xref="GI:9886711"
                     /translation="MEGDGGTPWALALLRTFDAGEFTGWEKVGSGGFGQVYKVRHVHW
                     KTWLAIKCSPSLHVDDRERMELLEEAKKMEMAKFRYILPVYGICREPVGLVMEYMETG
                     SLEKLLASEPLPWDLRFRIIHETAVGMNFLHCMAPPLLHLDLKPANILLDAHYHVKIS
                     DFGLAKCNGLSHSHDLSMDGLFGTIAYLPPERIREKSRLFDTKHDVYSFAIVIWGVLT
                     QKKPFADEKNILHIMVKVVKGHRPELPPVCRARPRACSHLIRLMQRCWQGDPRVRPTF
                     QEITSETEDLCEKPDDEVKETAHDLDVKSPPEPRSEVVPARLKRASAPTFDNDYSLSE
                     LLSQLDSGVSQAVEGPEELSRSSSESKLPSSGSGKRLSGVSSVDSAFSSRGSLSLSFE
                     REPSTSDLGTTDVQKKKLVDAIVSGDTSKLMKILQPQDVDLALDSGASLLHLAVEAGQ
                     EECAKWLLLNNANPNLSNRRGSTPLHMAVERRVRGVVELLLARKISVNAKDEDQWTAL
                     HFAAQNGDESSTRLLLEKNASVNEVDFEGRTPMHVACQHGQENIVRILLRRGVDVSLQ
                     GKDAWLPLHYAAWQGHLPIVKLLAKQPGVSVNAQTLDGRTPLHLAAQRGHYRVARILI
                     DLCSDVNVCSLLAQTPLHVAAETGHTSTARLLLHRGAGKKAVTSDGYTALHLAARNGH
                     LATVKLLVEEKADVLARGPLNQTALHLAAAHGHSEVVEELVSADVIDLFDEQGLSALH
                     LAAQGRHAQTVETLLRHGAHINLQSLKFQGGHGPAATLLRRSKT"
     polyA_signal    3838. .3843
                     /gene="dik"
BASE COUNT      807 a   1074 c   1178 g    820 t
ORIGIN


alignment_scores:
           Quality:  549.00          Length:     525
             Ratio:  1.913             Gaps:      25
 Percent Similarity:  54.667   Percent Identity: 32.762

alignment_block:
 US-09-445-223-1 x HSA278016   ..

 Align seg 1/1  to: HSA278016  from: 1  to: 3879

  27 GlyAlaSerGlyThrValSerSerAlaArgHisAlaAspTrpArgValGl 43
     |||:::   |||   |||   ::::::||||||:::   |||::::::
 139 GGCGGCTTCGGGCAGGTGTACAAGGTGCGCCATGTCCACTGGAAGACCTG 188

  43 nValAlaValLys.........HisLeuHisIleHisThrProLeuLeuA 57
      :::|||:::|||            ||||||:::               |
 189 GCTGGCCATCAAGTGCTCGCCCAGCCTGCACGTCGAC............G 226

  57 spSerGluArgLysAspValLeuArgGluAlaGluIleLeuHisLysAla 73
     ||   ||||||   ::::::|||:::|||||||:::  ::::::   |||
 227 ACAGGGAGCGCATGGAGCTTTTGGAAGAAGCCAAGAAGATGGAGATGGCC 276

  74 ArgPheSerTyrIlePheProIleLeuGlyIleCysAsnGluProGluPh 90
     :::|||   ||||||:::|||:::   |||||||||:::||||||
 277 AAGTTTCGCTACATCCTGCCTGTGTATGGCATCTGCCGCGAACCT..... 321

  90 eLeuGlyIleValThrGluTyrMetProAsnGlySerLeuAsnGluLeuL 107
      :::|||:::|||   |||||||||   :::|||||||||::::::||||
 322 .GTCGGCCTGGTCATGGAGTACATGGAGACGGGCTCCCTGGAAAAGCTGC 370

 107 euHisArgLysThrGluTyrProAspValAlaTrpProLeuArgPheArg 123
     ||      :::               :::   |||   ||||||||||||
 371 TGGCTTCGGAG............CCATTGCCATGGGATCTCCGGTTCCGA 408

 124 IleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisAsnMetThrPr 140
     |||:::||||||   |||:::|||:::|||:::||||||   |||:::||
 409 ATCATCCACGAGACGGCGGTGGGCATGAACTTCCTGCACTGCATGGCCCC 458

 140 oProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeuAspAsnG 157
     |||||||||||||   |||||||||     ||||||||||||||||   :
 459 GCCACTCCTGCACCTGGACCTCAAGCCCGCGAACATCCTGCTGGATGCCC 508

 157 luPheHisValLysIleAlaAspPheGlyLeuSerLysTrpArgMetMet 173
     :::::||||||||||||:::||||||||||||:::|||   :::   :::
 509 ACTACCACGTCAAGATTTCTGATTTTGGTCTGGCCAAGTGCAACGGGCTG 558
```

```
 174 SerLeuSerGlnSerArgSerSerLysSerAlaProGluGly......Gl 188
     |||   |||::::::   |||                :::|||      ||
 559 TCCCACTCGCATGACCTCAGCATG.............GATGGCCTGTTTGG 596

 188 yThrIleIleTyrMetProProGluAsnTyrGluProGlyGlnLysSerA 205
     |||||||   |||:::||||||||||:::   :::      :::|||||||
 597 CACAATCGCCTACCTCCCTCCAGAGCGCATCAGG......GAGAAGAGCC 640

 205 rgAla...SerIleLysHisAspIleTyrSerTyrAlaValIleThrTrp 220
     ||      :::   |||||||||:::||||||:::|||::::::   |||
 641 GGCTCTTCGACACCAAGCACGATGTATACAGCTTTGCGATCGTCATCTGG 690

 221 GluValLeuSerArgLysGlnProPheGluAspValThrAsnProLeuGl 237
        ||||||::::::|||:::||||||    |||      |||   |||::
 691 GGCGTGCTCACACAGAAGAAGCCGTTTGCAGATGAGAAGAACATCCTGCA 740

 237 nIleMetTyrSerValSerGlnGlyHisArgProValIleAsnGluGluS 254
     :||||||   :::|||   :::|||||||||||||               :
 741 CATCATGGTGAAGGTGGTGAAGGGCCACCGCCCC...............G 775

 254 erLeuProTyrAspIleProHisArgAlaArg...........MetIle 266
     ::||||||                |||   |||            :::|||
 776 AGCTGCCGCCCGTGTGCAGAGCCCGGCCGCGCGCCTGCAGCCACCTGATA 825

 267 SerLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerPh 283
        |||::::::      |||       :::|||      ||||||
 826 CGCCTCATGCAGCGGTGCTGGCAGGGGGATCCGCGAGTTAGGCCC..... 870

 283 eLeuLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGluI 300
                                           ||||||:::||||
 871 ......................................ACCTTCCAAGAAA 883

 300 leThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSer 316
     |||||
 884 TTACT.............................................. 888

 317 ValSerSerAlaIleHisLeuCysAspLys..............LysLy 328
        |||::::::      ||||||:::|||                |||::
 889 ...TCTGAAACCGAGGACCTGTGTGAAAAGCCTGATGACGAAGTGAAAGA 935

 328 sMetGluLeuSerLeuAsnIleProValAsnHisGlyProGlnGluGluS 345
     :         :::|||::::::                |||::::::|||
 936 AACTGCTCATGATCTGGACGTGAAAAGCCCCCCGGAGCCCAGGAGCGAGG 985

 345 er.CysGlySerSerGlnLeuHisGluAsnSerGlySerProGluThrSe 361
        |||            |||:::    :::||||||
 986 TGGTGC............CTGCGAGGCTCAAGCGGG.............. 1009

 361 rArgSerLeuProAlaProGlnAspAsnAspPheLeuSerArgLysAlaG 378
            ||||||   |||:::             ::::::      |||:
1010 ....CCTCTGCCCCCACCTTCG............ATAACGACTACAGCCT 1043

 378 lnAspCysTyrPheMetLysLeuHisHisCysProGlyAsnHisSerTrp 394
     ::      :::                            :::|||||||||
1044 CTCCGAGCTTC.........................TCTCACAGCTGG 1066

 395 AspSerThrIleSerGlySerGlnArgAlaAlaPheCysAspHisLysTh 411
              :::         :::||||||
1067 ACTCTGGAGTTTCCCAGGCTGTCGAGGGCC..................... 1096

 411 rThrProCysSerSerAlaIle......IleAsnPro..LeuSerThrAl 425
         |||   |||||||||          ::::::|||  |||   :::::
1097 ....CCGAGGAGCTCAGCCGCAGCTCCTCTGAGTCCAAGCTGCCATCGTC 1142

 425 aGlyAsnSerGluArgLeuGlnProGlyIleAlaGlnGlnTrpIleGlnS 442
     :|||:::::::::||||||:::    |||::::::
1143 CGGCAGTGGGAAGAGGCTCTCG...GGGGTGTCC................ 1173

 442 erLysArgGluAspIleValAsnGlnMetThrGluAlaCysLeuAsnGln 458

1173 .................................................. 1173
```

```
      459 SerLeuAspAlaLeuLeuSerArgAspLeuIleMetLysGluAspTyrGl 475
          |||:::|||:::   :::|||             ::::::   ::::::||
     1174 TCGGTGGACTCCGCCTTCTCTTCCAGAGGATCACTGTCGCTGTCCTTTGA 1223

      475 uLeuValSerThrLysProThrArgThrSerLysValArgGlnLeuLeuA 492
          |            :::|||:::    ::::::               |||
     1224 G........CGGGAACCTTCAACCAGCGAT...............CTGG 1249

      492 spThrThrAspIleGlnGlyGluGluPheAlaLysValIleValGlnLys 508
            |||||||||:::|||    ::::::::::::   :::||||||:::
     1250 GTACCACAGACGTCCAGAAGAAGAAGCTTGTGGATGCCATCGTGTCCGGG 1299

      509 LeuLysAspAsnLysGlnMetGlyLeuGlnProTyrProGluIleLeuVa 525
                ::::::   :::      |||||||||             |||::
     1300 GACACCAGCAAACTGATGAAGATCCTGCAGCCGCAGGACGTGGACCTGGC 1349

      525 lValSerArgSerProSerLeu 532
          :::::::   :::   ||||||
     1350 ACTGGACAGCGGTGCCAGCCTG 1371

seq_name: gb_rol:AF178953

seq_documentation_block:
LOCUS       AF178953     1833 bp    mRNA            ROD       19-OCT-1999
DEFINITION  Mus musculus receptor interacting protein 3 (Rip3) mRNA, complete
            cds.
ACCESSION   AF178953
VERSION     AF178953.1  GI:6063100
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1833)
  AUTHORS   Pazdernik,N.J., Donner,D.B., Goebl,M.G. and Harrington,M.A.
  TITLE     Mouse Interacting Protein 3 Does Not Contain a Caspase-Recruiting
            or a Death Domain but Induces Apoptosis and Activates NF-kappaB
  JOURNAL   Mol. Cell. Biol. 19 (1999) In press
REFERENCE   2  (bases 1 to 1833)
  AUTHORS   Pazdernik,N.J., Donner,D.B., Goebl,M.G. and Harrington,M.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-AUG-1999) Walther Oncology Center, Indiana University
            School of Medicine, 1044 West Walnut Street, Indianapolis, IN
            46202, USA
FEATURES             Location/Qualifiers
     source          1. .1833
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     gene            1. .1833
                     /gene="Rip3"
     misc_feature    151. .894
                     /gene="Rip3"
                     /note="Region: N-terminal protein kinase domain"
     CDS             151. .1611
                     /gene="Rip3"
                     /codon_start=1
                     /product="receptor interacting protein 3"
                     /protein_id="AAF03133.1"
                     /db_xref="GI:6063101"
                     /translation="MSSVKLWPTGASAVPLVSREELKKLEFVGKGGFGVVFRAHHRTW
                     NHDVAVKIVNSKKISWEVKAMVNLRNENVLLLLGVTEDLQWDFVSGQALVTRFMENGS
                     LAGLLQPECPRPWPLLCRLLQEVVLGMCYLHSLDPPLLHRDLKPSNILLDPELHAKLA
                     DFGLSTFQGGSQSGSGSGSGSRDSGGTLAYLDPELLFKVNLKASKASDVYSFGILVWA
                     VLAGREAELVDKTSLIRETVCDRQSRPPLTELPPGSPETPGLEKLKELMIHCWGSQSE
                     NRPSFQDCEPKTNEVYNLVKDKVDAAVSEVKHYLSQHRSSGRNLSAREPSQRGTEMDC
                     PRETMVSKMLDRLHLEEPSGPVPGKCPERQAQDTSVGPATPARTSSDPVAGTPQIPHT
                     LPFRGTTPGPVFTETPGPHPQRNQGDGRHGTPWYPWTPPNPMTGPPALVFNNCSEVQI
                     GNYNSLVAPPRTTASSSAKYDQAQFGRGRGWQPFHK"
BASE COUNT      466 a    497 c    477 g    393 t
ORIGIN


alignment_scores:
```

```
         Quality:  402.00           Length:    534
           Ratio:  1.529              Gaps:     19
Percent Similarity:  49.251   Percent Identity: 27.154

alignment_block:
 US-09-445-223-1 x AF178953   ..

 Align seg 1/1  to: AF178953  from: 1  to: 1833

   8 SerAlaLeuProThrIleProTyrHisLysLeuAlaAspLeuArgTyrLe 24
     ||||||:::|||   :::      :::::::|||     |||::::::::
 184 TCAGCGGTTCCTCTGGTGAGCCGTGAAGAACTGAAGAAGCTGGAGTTTGT 233

  24 uSerArgGlyAlaSerGlyThrValSerSerAlaArgHisAlaAspTrpA 41
     :::::::|||:::   |||:::|||      |||:::|||      |||:
 234 GGGCAAAGGAGGGTTCGGAGTCGTGTTCCGGGCACACCACAGAACATGGA 283

  41 rgValGlnValAlaValLysHisLeuHisIleHisThrProLeuLeuAsp 57
     ::   :::|||||||||||||  ::::::
 284 ACCATGATGTAGCAGTCAAGATCGTGAAC..................... 312

  58 SerGluArgLysAspValLeuArgGluAlaGluIleLeuHisLysAlaAr 74
              |||   :::      |||::::::   :::   :::   ||
 313 ......TCGAAGAAGATATCCTGGGAGGTGAAGGCTATGGTTAATCTTCG 356

  74 gPheSerTyrIlePheProIleLeuGlyIleCysAsnGluProGluPhe. 90
     |   :::   ::::::   :::|||||||:::  ::::::   ::::::
 357 TAATGAGAACGTTCTGCTCCTGCTGGGGGTCACTGAGGACCTCCAGTGGG 406

  91 ..............LeuGlyIleValThrGluTyrMetProAsnGlySer 102
                      ::::::||||||::::::|||   |||||||||
 407 ACTTCGTGTCCGGGCAGGCTCTGGTGACAAGATTCATGGAGAATGGCTCC 456

 103 LeuAsnGluLeuLeuHisArgLysThrGluTyrProAspValAlaTrpPr 119
     |||      ||||||:::   :::         |||         |||||
 457 CTCGCAGGGCTGCTGCAACCCGAGTGCCCTCGGCCC.........TGGCC 497

 119 oLeuArgPheArgIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuH 136
     ||||       |||:::|||:::|||::::::||||||:::  |||||||
 498 ACTCCTCTGTCGCCTGCTGCAGGAAGTGGTGCTGGGGATGTGCTACCTAC 547

 136 isAsnMetThrProProLeuLeuHisHisAspLeuLysThrGlnAsnIle 152
     ||::::::    |||||||||||||||:::|||||||||  :::||||||
 548 ACAGCTTGGACCCTCCGCTCCTGCACCGGGACCTCAAGCCCTCTAACATT 597

 153 LeuLeuAspAsnGluPheHisValLysIleAlaAspPheGlyLeuSerLy 169
     |||||||||   |||:::|||:::|||:::||||||||||||||||||
 598 CTGCTGGATCCAGAGCTCCACGCCAAGCTAGCAGATTTTGGCCTGTCCAC 647

 169 sTrpArgMetMetSerLeuSerGlnSerArgSer...SerLysSerAlaP 185
      ::::::      |||   |||   |||   |||   |||   |||
 648 GTTTCAGGGAGGGTCCCAGTCAGGGTCAGGATCAGGATCAGGATCCAGGG 697

 185 roGluGlyGlyThrIleIleTyrMetProProGlu............... 196
       :::|||||||||:::   |||:::   ||||||
 698 ACTCTGGGGGCACCCTAGCGTACTTGGACCCAGAGCTGTTATTTAAAGTC 747

 197 AsnTyrGluProGlyGlnLysSerArgAlaSerIleLysHisAspIleTy 213
     |||                     :::||||||         |||:::||
 748 AACCTG..................AAGGCTTCTAAAGCGAGTGATGTCTA 779

 213 rSerTyrAlaValIleThrTrpGluValLeuSer...ArgLysGlnProP 229
     ||||:::::::::::::::|||   ||||||:::   |||:::      :
 780 CAGCTTTGGGATCCTCGTGTGGGCAGTGCTGGCTGGCAGAGAAGCTGAGT 829

 229 heGluAspValThrAsnProLeuGlnIleMetTyrSerValSerGlnGly 245
     ::   |||   |||:::   ::::::                  |||:::
 830 TGGTAGACAAGACTTCACTAATCCGGGAAACAGTGTGTGACAGGCAGAGT 879

 246 HisArgProValIleAsnGluGluSerLeuProTyrAspIleProHisAr 262
     :::   |||:::   :::                  :::   |||
 880 CGTCCTCCACTGACAGAGCTGCCTCCAGGTAGCCCTGAGACTCCCGGCTT 929

 262 gAlaArgMetIleSerLeuIleGluSerGlyTrpAlaGlnAsnProAspG 279
         ::::::   :::|||:::         |||:::::::::   ::::
 930 GGAAAAACTGAAGGAGTTAATGATTCATTGCTGGGGTTCCCAGTCCGAAA 979

 279 luArgProSerPheLeuLysCysLeuIleGluLeuGluProValLeuArg 295
     ::||||||||||||      |||            ||||||      :::
 980 ACAGGCCATCCTTCCAGGACTGC............GAACCA......AAA 1011

 296 ThrPheGluGluIleThrPheLeuGlu...............AlaValIl 307
     |||   |||      ::::::::::::               ||||||
1012 ACCAATGAAGTTTACAATCTGGTAAAGGACAAGGTAGATGCTGCTGTCTC 1061

 307 eGlnLeuLysLysThrLysLeuGlnSerValSerSerAlaIleHisLeu. 323
      ::::::|||            |||      ||||||:::   :::|||
1062 CGAGGTAAAGCATTATCTGTCTCAGCACAGAAGCAGTGGCAGAAACTTGT 1111

 324 ......................................CysAspLysLys 327
                                           |||   ::::::
1112 CTGCCAGAGAGCCAAGCCAAAGAGGCACAGAAATGGATTGCCCGAGGGAA 1161

 328 ............LysMetGluLeuSerLeuAsnIleProValAsnHisGl 340
                 ||||||         |||::::::            ||
1162 ACCATGGTTTCTAAAATGCTGGACCGCCTGCATTTGGAGGAACCCTCCGG 1211

 340 yProGlnGluGluSerCysGlySerSerGlnLeuHisGluAsnSerGlyS 357
     ||||         :::|||   :::   |||   :::::::::|||   :
1212 ACCAGTTCCTGGAAAATGTCCTGAGAGGCAAGCACAGGACACATCAGTTG 1261

 357 erProGluThrSerArgSerLeuProAlaProGlnAspAsnAspPheLeu 373
     ::|||   |||            ||||||      ::::::|||
1262 GGCCTGCCACA............CCAGCAAGGACATCTTCTGAC...... 1293

 374 SerArgLysAlaGlnAspCysTyrPheMetLysLeuHisHisCysPro.. 389
                                                  |||
1294 .............................................CCCGT 1298

 390 GlyAsnHisSerTrpAspSerThrIleSerGlySerGlnArg........ 403
     |||   ||||||   |||||||||            ||||||
1299 GGCTGGCACTCCTCAGATTCCACATACTTTACCCTTCAGAGGCACAACAC 1348

 404 .......AlaAlaPheCysAspHisLysThrThrProCysSerSerAlaI 418
            :::   :::   |||   :::::::::|||
1349 CTGGGCCAGTCTTTACTGAGACTCCCGGTCCTCACCC............. 1385

 418 leIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeuGlnProGly 434
                                                 |||
1386 ............................................CCAAAG 1391

 435 IleAlaGlnGlnTrpIleGlnSerLysArgGluAspIleValAsnGlnMe 451
        :::   :::|||
1392 GAATCAGGGAGATGG................................... 1406

 451 tThrGluAlaCysLeuAsnGlnSerLeuAspAlaLeuLeuSerArgAspL 468

1406 .................................................. 1406

 468 euIleMetLysGluAspTyrGluLeuValSerThrLysProThrArgThr 484
             |||      ::::::|||||||||      ||||||::::::
1407 ........AAGACACGGCACTCCTTGGTATCCCTGGACCCCACCGAATCC 1448

 485 ..................SerLysValArgGlnLeuLeuAspThrThrAs 495
                       |||:::::::::||||||:::   ::::::||
1449 AATGACAGGGCCACCGGCTCTCGTCTTCAACAACTGTTCTGAAGTGCAGA 1498

 495 p 495
     |
1499 T 1499

seq_name: gb_pat2:I68122

seq_documentation_block:
LOCUS       I68122       2268 bp    DNA             PAT       04-FEB-1998
DEFINITION  Sequence 14 from patent US 5674734.
```

```
ACCESSION   I68122
VERSION     I68122.1  GI:2830244
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2268)
  AUTHORS   Leder,P., Seed,B., Stanger,B.Z., Lee,T. and Kim,E.
  TITLE     Cell death protein
  JOURNAL   Patent: US 5674734-A 14 07-OCT-1997;
FEATURES             Location/Qualifiers
     source          1..2268
                     /organism="unknown"
BASE COUNT      719 a    491 c    576 g    482 t
ORIGIN


alignment_scores:
           Quality:  397.00             Length:    609
             Ratio:  1.272                Gaps:     25
 Percent Similarity:  51.232   Percent Identity: 25.452

alignment_block:
 US-09-445-223-1 x I68122   ..

 Align seg 1/1  to: I68122  from: 1  to: 2268

   24 LeuSerArgGlyAlaSerGlyThrValSerSerAlaArgHisAlaAspTr 40
      |||:::   |||:::   |||   ||||||   :::   |||   :::
  118 CTAGACAGCGGAGGCTTCGGGAAGGTGTCCTTGTGTTACCACAGAAGCCA 167

   40 pArgValGlnValAlaValLysHisLeuHisIleHisThrProLeuLeuA 57
                :::      |||         ::::::|||
  168 TGGATTTGTCATCCTGAAAAAA.........GTATACACAGGGCCCAACC 208

   57 spSerGluArgLysAspVal...LeuArgGluAlaGluIleLeuHisLys 72
        :::|||   ::::::|||   |||:::|||::::::::::::|||:::
  209 GCGCTGAGTACAATGAGGTTCTCTTGGAAGAGGGGAAGATGATGCACAGA 258

   73 AlaArgPheSerTyrIlePheProIleLeuGlyIleCysAsnGluProGl 89
         |||   |||   :::      :::|||||||||      |||
  259 CTGAGACACAGTCGAGTGGTGAAGCTACTGGGCATCATCATAGAAGAAGG 308

   89 uPheLeuGlyIleValThrGluTyrMetProAsnGlySerLeuAsnGluL 106
             ::::::|||   |||||||||   :::|||:::|||   ::::
  309 GAACTATTCGCTGGTGATGGAGTACATGGAGAAGGGCAACCTGATGCACG 358

  106 euLeuHisArgLysThrGluTyrProAspValAlaTrpProLeuArgPhe 122
      ::|||      ||||||:::      ||||||         |||:::
  359 TGCTA......AAGACCCAGATA...GATGTCCCACTTTCATTGAAAGGA 399

  123 ArgIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisAsnMetTh 139
      ||||||:::   |||         |||:::   |||||||||:::
  400 AGGATAATCGTGGAGGCCATAGAAGGCATGTGCTACTTACATGAC..... 444

  139 rProProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeuAspA 156
             ::::::|||   |||||||||   :::|||||||||:::|||:
  445 .AAAGGTGTGATACACAAGGACCTGAAGCCTGAGAATATCCTCGTTGATC 493

  156 snGluPheHisValLysIleAlaAspPheGlyLeuSerLysTrpArgMet 172
      :::::||||||:::||||||||||||:::|||:::::::::::::::
  494 GTGACTTTCACATTAAGATAGCCGATCTTGGTGTGGCTTCCTTTAAGACA 543

  173 MetSer...LeuSerGlnSerArgSerSerLysSerAlaPro........ 185
         |||   |||::::::::::::::::::|||:::
  544 TGGAGCAAACTGACTAAGGAGAAAGACAACAAGCAGAAAGAAGTGAGCAG 593

  186 ................GluGlyGlyThrIleIleTyrMetProProGlu. 196
                      :::|||||||||:::   ||||||   ||||||
  594 CACCACTAAGAAGAACAATGGTGGTACCCTTTACTACATGGCACCCGAAC 643

  197 ..............AsnTyrGluProGlyGlnLysSerArgAlaSerIle 208
                    |||   :::|||   :::||||||
  644 ACCTGAATGACATCAATGCAAAGCCCACGGAGAAGTCG............ 681

  209 LysHisAspIleTyrSerTyrAlaValIleThrTrpGluValLeuSerAr 225
            |||:::||||||::::::::::::   |||   :::::::::::
  682 ......GACGTGTACAGCTTTGGCATTGTCCTTTGGGCAATATTTGCAAA 725

  225 gLysGlnProPheGluAspValThrAsnProLeuGlnIleMetTyrSerV 242
      :|||:::|||:::|||:::|||            |||::::::      :
  726 AAAGGAGCCCTATGAGAATGTCATCTGTACTGAGCAGTTCGTGATCTGCA 775

  242 alSerGlnGlyHisArgProValIleAsnGlu......GluSerLeuPro 256
      ::::::::|||:::||||||   ::::::|||      |||      |||
  776 TAAAATCTGGGAACAGGCCAAATGTAGAGGAAATCCTTGAGTACTGTCCA 825

  257 TyrAspIleProHisArgAlaArgMetIleSerLeuIleGluSerGlyTr 273
         :::|||                  |||||||||:::|||      ||
  826 AGGGAGATC.................ATCAGCCTCATGGAGCGGTGCTG 857

  273 pAlaGlnAsnProAspGluArgProSerPheLeuLysCysLeuIleGluL 290
      |         |||::::::||||||:::||||||            |||:
  858 GCAGGCGATCCCAGAAGACAGGCCAACATTTCTTGGCATTGAAGAAGAAT 907

  290 euGluPro...ValLeuArgThrPheGluGluIleThrPheLeuGluAla 305
      :::::|||      |||      |||||||||   :::      |||
  908 TTAGGCCTTTTTACTTAAGTCATTTTGAAGAATATGTA...GAAGAGGAT 954

  306 ValIleGlnLeuLysLys..................ThrLysLeuGlnSe 316
      |||   :::|||||||||                        ||||||
  955 GTGGCAAGTTTAAAGAAAGAGTATCCAGATCAAAGCCCAGTGCTGCAGAG 1004

  316 rValSerSerAlaIleHisLeuCys......................... 324
       :::   |||      |||   |||
 1005 AATGTTTTCACTGCAGCATGACTGTGTACCCTTACCTCCGAGCAGGTCAA 1054

  325 ..AspLysLysLysMetGluLeuSerLeuAsnIleProValAsnHisGly 340
        :::::::::      :::|||      :::      ::::::   |||
 1055 ATTCAGAACAACCTGGATCGCTGCACAGTTCCCAGGGGCTCCAGATGGGT 1104

  341 ProGlnGluGluSerCysGlySerSerGln............LeuHisGl 353
      |||   |||||||||      ||||||:::                  ||
 1105 CCTGTGGAGGAGTCCTGGTTTTCTTCCTCCCCAGAGTACCCACAGGACGA 1154

  353 uAsnSerGlySerProGluThr............................ 360
      ||||:::   |||   ::::::
 1155 GAATGATCGCAGTGTGCAGGCTAAGCTGCAAGAGGAAGCCAGCTATCATG 1204

  361 ....................SerArgSerLeuProAlaProGlnAspAsn 370
                          :::::::::   |||   |||::::::|||
 1205 CTTTTGGAATATTTGCAGAGAAACAGACAAAACCGCAGCCAAGGCAGAAT 1254

  371 AspPheLeuSerArgLysAlaGlnAspCysTyrPheMetLysLeuHisHi 387
      :::      :::|||:::   :::               ::::::   ||
 1255 GAGGCTTACAACAGAGAGGAGGAA......AGGAAACGAAGGGTCTCTCA 1298

  387 sCysPro........................................... 389
      |   |||
 1299 TGACCCCTTTGCACAGCAGAGAGCTCGTGAGAATATTAAGAGTGCAGGAG 1348

  390 ..GlyAsnHisSerTrpAspSerThrIleSerGlySerGlnArgAlaAla 405
           :::||||||      ||||||   |||         |||:::
 1349 CAAGAGGTCATTCTGATCCCAGCACAACGAGT.........CGTGGAATT 1389

  406 PheCysAspHisLysThrThrProCysSerSerAlaIleIleAsn..... 420
            ::::::   :::   |||:::::::::::::::   |||
 1390 GCAGTGCAACAGCTGTCATGGCCAGCCACCCAAACAGTTTGGAACAATGG 1439

  421 ................ProLeuSerThrAlaGlyAsnSerGluArgLeuG 432
                         ::::::|||:::|||::::::
 1440 ATTGTATAATCAGCATGGATTTGGAACTACAGGTACAGGAGTTTGGTATC 1489

  432 lnProGlyIleAlaGlnGlnTrpIleGlnSerLys............... 443
        |||:::::::::|||::::::         |||
 1490 CGCCAAATCTAAGCCAAATGTATAGTACTTATAAAACTCCAGTGCCAGAG 1539
```

```
 443 .................................................. 443

1540 ACCAACATACCGGGAAGCACACCCACCATGCCATACTTCTCTGGGCCAGT 1589

 444 .ArgGluAspIleVal.AsnGlnMet........................ 451
         :::|||:::::: |||   :::
1590 AGCAGATGACCTCATAAAATATACTATATTCAATAGTTCTGGTATTCAGA 1639

 452 ThrGluAlaCysLeuAsnGlnSerLeuAspAlaLeuLeuSerArgAspLe 468
        |||:::   :::         ||||||   :::   :::::::::
1640 TTGGAAACCACAATTATATGGATGTTGGACTGAATTCACAACCACCAAAC 1689

 468 uIleMetLysGluAspTyrGluLeuValSerThrLysProThrArgThrS 485
      |||:::   :::      :::|||   :::|||||||||:::      :
1690 AATACTTGCAAAGAAGAGTCGACTTCCAGACACCAAGCCATCTTTGATAA 1739

 485 erLysValArgGlnLeuLeuAspThrThrAspIleGlnGlyGlu...... 499
     ::   :::      |||::::::|||      ::::::|||:::
1740 CACCACTAGTCTGACTGATGAACACCTGAACCCTATCAGGGAAAACCTGG 1789

 500 ...........................GluPheAlaLysValIleValGl 507
                                ::::::|||::::::   :::::
1790 GAAGGCAGTGGAAAAACTGTGCCCGCAAGCTGGGCTTCACTGAGTCTCAG 1839

 507 nLysLeuLysAspAsnLysGlnMet 515
     :::::::|||::::::      |||
1840 ATCGATGAAATCGACCATGACTATG 1864

seq_name: gb_rol:MMU25995

seq_documentation_block:
LOCUS       MMU25995     2268 bp    mRNA            ROD       24-MAY-1995
DEFINITION  Mus musculus cell death protein (RIP) mRNA, complete cds.
ACCESSION   U25995
VERSION     U25995.1  GI:829618
KEYWORDS    Fas; TNF receptor.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2268)
  AUTHORS   Stanger,B.Z., Leder,P., Lee,T.H., Kim,E. and Seed,B.
  TITLE     RIP: a novel protein containing a death domain that interacts with
            Fas/APO-1 (CD95) in yeast and causes cell death
  JOURNAL   Cell 81 (4), 513-523 (1995)
  MEDLINE   95277838
REFERENCE   2  (bases 1 to 2268)
  AUTHORS   Stanger,B.Z.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAY-1995) Ben Z. Stanger, Genetics, Harvard Medical
            School, 200 Longwood Avenue, Boston, MA 02115, USA
FEATURES             Location/Qualifiers
     source          1..2268
                     /organism="Mus musculus"
                     /strain="(C57Bl/6 X CBA)F1"
                     /db_xref="taxon:10090"
                     /tissue_type="thymus"
                     /clone="III2"
     gene            52..2022
                     /gene="RIP"
     CDS             52..2022
                     /gene="RIP"
                     /codon_start=1
                     /function="Fas interacting protein; cell death"
                     /product="RIP"
                     /protein_id="AAB60487.1"
                     /db_xref="GI:829619"
                     /translation="MQPDMSLDNIKMASSDLLEKTDLDSGGFGKVSLCYHRSHGFVIL
                     KKVYTGPNRAEYNEVLLEEGKMMHRLRHSRVVKLLGIIIEEGNYSLVMEYMEKGNLMH
                     VLKTQIDVPLSLKGRIIVEAIEGMCYLHDKGVIHKDLKPENILVDRDFHIKIADLGVA
                     SFKTWSKLTKEKDNKQKEVSSTTKKNNGGTLYYMAPEHLNDINAKPTEKSDVYSFGIV
                     LWAIFAKKEPYENVICTEQFVICIKSGNRPNVEEILEYCPREIISLMERCWQAIPEDR
```

```
                     PTFLGIEEEFRPFYLSHFEEYVEEDVASLKKEYPDQSPVLQRMFSLQHDCVPLPPSRS
                     NSEQPGSLHSSQGLQMGPVEESWFSSSPEYPQDENDRSVQAKLQEEASYHAFGIFAEK
                     QTKPQPRQNEAYNREEERKRRVSHDPFAQQRARENIKSAGARGHSDPSTTSRGIAVQQ
                     LSWPATQTVWNNGLYNQHGFGTTGTGVWYPPNLSQMYSTYKTPVPETNIPGSTPTMPY
                     FSGPVADDLIKYTIFNSSGIQIGNHNYMDVGLNSQPPNNTCKEESTSRHQAIFDNTTS
                     LTDEHLNPIRENLGRQWKNCARKLGFTESQIDEIDHDYERDGLKEKVYQMLQKWLMRE
                     GTKGATVGKLAQALHQCCRIDLLNHLIRASQS"
     variation       1469
                     /gene="RIP"
                     /note="apparent polymorphism resulting in a Thr vs Ile
                     variation"
                     /replace="T"
     variation       1623
                     /gene="RIP"
                     /note="silent polymorphism"
                     /replace="C"
BASE COUNT      719 a    490 c    577 g    482 t
ORIGIN


alignment_scores:
         Quality:   397.00            Length:    609
           Ratio:   1.272               Gaps:     25
 Percent Similarity:  51.232   Percent Identity:  25.452

alignment_block:
 US-09-445-223-1 x MMU25995   ..

 Align seg 1/1  to: MMU25995  from: 1  to: 2268

  24 LeuSerArgGlyAlaSerGlyThrValSerSerAlaArgHisAlaAspTr 40
     |||:::   |||:::   |||   ||||||   :::   |||   :::
 118 CTAGACAGCGGAGGCTTCGGGAAGGTGTCCTTGTGTTACCACAGAAGCCA 167

  40 pArgValGlnValAlaValLysHisLeuHisIleHisThrProLeuLeuA 57
               :::      |||         ::::::|||
 168 TGGATTTGTCATCCTGAAAAAA.........GTATACACAGGGCCCAACC 208

  57 spSerGluArgLysAspVal...LeuArgGluAlaGluIleLeuHisLys 72
       :::|||   ::::::|||   |||:::|||::::::::::::|||:::
 209 GCGCTGAGTACAATGAGGTTCTCTTGGAAGAGGGGAAGATGATGCACAGA 258

  73 AlaArgPheSerTyrIlePheProIleLeuGlyIleCysAsnGluProGl 89
        |||   |||   :::      :::|||||||||      |||
 259 CTGAGACACAGTCGAGTGGTGAAGCTACTGGGCATCATCATAGAAGAAGG 308

  89 uPheLeuGlyIleValThrGluTyrMetProAsnGlySerLeuAsnGluL 106
            ::::::|||   |||||||||   :::|||:::|||   ::::
 309 GAACTATTCGCTGGTGATGGAGTACATGGAGAAGGGCAACCTGATGCACG 358

 106 euLeuHisArgLysThrGluTyrProAspValAlaTrpProLeuArgPhe 122
     ::|||      ||||||:::      ||||||         |||:::
 359 TGCTA......AAGACCCAGATA...GATGTCCCACTTTCATTGAAAGGA 399

 123 ArgIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisAsnMetTh 139
     ||||||:::   |||         |||:::   |||||||||:::
 400 AGGATAATCGTGGAGGCCATAGAAGGCATGTGCTACTTACATGAC..... 444

 139 rProProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeuAspA 156
            ::::::|||   |||||||||   :::|||||||||:::|||:
 445 .AAAGGTGTGATACACAAGGACCTGAAGCCTGAGAATATCCTCGTTGATC 493

 156 snGluPheHisValLysIleAlaAspPheGlyLeuSerLysTrpArgMet 172
     :::::||||||:::||||||||||||:::|||:::::::::::::::
 494 GTGACTTTCACATTAAGATAGCCGATCTTGGTGTGGCTTCCTTTAAGACA 543

 173 MetSer...LeuSerGlnSerArgSerSerLysSerAlaPro........ 185
        |||   |||::::::::::::::::::|||:::
 544 TGGAGCAAACTGACTAAGGAGAAAGACAACAAGCAGAAAGAAGTGAGCAG 593

 186 ................GluGlyGlyThrIleIleTyrMetProProGlu. 196
                     :::|||||||||:::   ||||||   ||||||
 594 CACCACTAAGAAGAACAATGGTGGTACCCTTTACTACATGGCACCCGAAC 643
```

```
 197 ..............AsnTyrGluProGlyGlnLysSerArgAlaSerIle 208
                   |||   :::|||   :::||||||
 644 ACCTGAATGACATCAATGCAAAGCCCACGGAGAAGTCG............ 681

 209 LysHisAspIleTyrSerTyrAlaValIleThrTrpGluValLeuSerAr 225
           |||:::||||||::::::::::::   |||   :::::::::::
 682 ......GACGTGTACAGCTTTGGCATTGTCCTTTGGGCAATATTTGCAAA 725

 225 gLysGlnProPheGluAspValThrAsnProLeuGlnIleMetTyrSerV 242
     :|||:::|||:::|||:::|||            |||::::::      :
 726 AAAGGAGCCCTATGAGAATGTCATCTGTACTGAGCAGTTCGTGATCTGCA 775

 242 alSerGlnGlyHisArgProValIleAsnGlu......GluSerLeuPro 256
     ::::::::|||:::||||||   ::::::|||      |||      |||
 776 TAAAATCTGGGAACAGGCCAAATGTAGAGGAAATCCTTGAGTACTGTCCA 825

 257 TyrAspIleProHisArgAlaArgMetIleSerLeuIleGluSerGlyTr 273
        :::|||                  |||||||||:::|||      ||
 826 AGGGAGATC..................ATCAGCCTCATGGAGCGGTGCTG 857

 273 pAlaGlnAsnProAspGluArgProSerPheLeuLysCysLeuIleGluL 290
     |         |||::::::||||||:::||||||            |||:
 858 GCAGGCGATCCCAGAAGACAGGCCAACATTTCTTGGCATTGAAGAAGAAT 907

 290 euGluPro...ValLeuArgThrPheGluGluIleThrPheLeuGluAla 305
     :::::|||      |||      |||||||||   :::      |||
 908 TTAGGCCTTTTTACTTAAGTCATTTTGAAGAATATGTA...GAAGAGGAT 954

 306 ValIleGlnLeuLysLys..................ThrLysLeuGlnSe 316
     |||   :::|||||||||                        ||||||
 955 GTGGCAAGTTTAAAGAAAGAGTATCCAGATCAAAGCCCAGTGCTGCAGAG 1004

 316 rValSerSerAlaIleHisLeuCys......................... 324
      :::   |||      |||   |||
1005 AATGTTTTCACTGCAGCATGACTGTGTACCCTTACCTCCGAGCAGGTCAA 1054

 325 ..AspLysLysLysMetGluLeuSerLeuAsnIleProValAsnHisGly 340
       :::::::::      :::|||      :::      ::::::   |||
1055 ATTCAGAACAACCTGGATCGCTGCACAGTTCCCAGGGGCTCCAGATGGGT 1104

 341 ProGlnGluGluSerCysGlySerSerGln............LeuHisGl 353
     |||   |||||||||      ||||||:::                  ||
1105 CCTGTGGAGGAGTCCTGGTTTCTTCCTCCCCAGAGTACCCACAGGACGA 1154

 353 uAsnSerGlySerProGluThr............................ 360
     ||||:::   |||   ::::::
1155 GAATGATCGCAGTGTGCAGGCTAAGCTGCAAGAGGAAGCCAGCTATCATG 1204

 361 ....................SerArgSerLeuProAlaProGlnAspAsn 370
                         :::::::::   |||   |||::::::|||
1205 CTTTTGGAATATTTGCAGAGAAACAGACAAAACCGCAGCCAAGGCAGAAT 1254

 371 AspPheLeuSerArgLysAlaGlnAspCysTyrPheMetLysLeuHisHi 387
     :::      :::|||:::   :::               ::::::   ||
1255 GAGGCTTACAACAGAGAGGAGGAA......AGGAAACGAAGGGTCTCTCA 1298

 387 sCysPro........................................... 389
     |   |||
1299 TGACCCCTTTGCACAGCAGAGAGCTCGTGAGAATATTAAGAGTGCAGGAG 1348

 390 ..GlyAsnHisSerTrpAspSerThrIleSerGlySerGlnArgAlaAla 405
          :::||||||      ||||||   |||         |||:::
1349 CAAGAGGTCATTCTGATCCCAGCACAACGAGT.........CGTGGAATT 1389

 406 PheCysAspHisLysThrThrProCysSerSerAlaIleIleAsn..... 420
           ::::::   :::   |||:::::::::::::::   |||
1390 GCAGTGCAACAGCTGTCATGGCCAGCCACCCAAACAGTTTGGAACAATGG 1439

 421 ................ProLeuSerThrAlaGlyAsnSerGluArgLeuG 432
                        ::::::|||:::|||::::::
1440 ATTGTATAATCAGCATGGATTTGGAACTACAGGTACAGGAGTTTGGTATC 1489
```

```
 432 lnProGlyIleAlaGlnGlnTrpIleGlnSerLys............... 443
       |||:::::::::|||::::::         |||
1490 CGCCAAATCTAAGCCAAATGTATAGTACTTATAAAACTCCAGTGCCAGAG 1539

 443 .................................................. 443

1540 ACCAACATACCGGGAAGCACACCCACCATGCCATACTTCTCTGGCCAGT 1589

 444 .ArgGluAspIleVal.AsnGlnMet........................ 451
         :::|||:::::: |||   :::
1590 AGCAGATGACCTCATAAAATATACTATATTCAATAGTTCTGGTATTCAGA 1639

 452 ThrGluAlaCysLeuAsnGlnSerLeuAspAlaLeuLeuSerArgAspLe 468
        |||:::   :::         ||||||   :::   :::::::::
1640 TTGGAAACCACAATTATATGGATGTTGGACTGAATTCACAACCACCAAAC 1689

 468 uIleMetLysGluAspTyrGluLeuValSerThrLysProThrArgThrS 485
      |||:::   :::      :::|||   :::|||||||||:::      :
1690 AATACTTGCAAAGAAGAGTCGACTTCCAGACACCAAGCCATCTTTGATAA 1739

 485 erLysValArgGlnLeuLeuAspThrThrAspIleGlnGlyGlu...... 499
     ::   :::      |||::::::|||      ::::::|||:::
1740 CACCACTAGTCTGACTGATGAACACCTGAACCCTATCAGGGAAAACCTGG 1789

 500 ...........................GluPheAlaLysValIleValGl 507
                                ::::::|||::::::   :::::
1790 GAAGGCAGTGGAAAAACTGTGCCCGCAAGCTGGGCTTCACTGAGTCTCAG 1839

 507 nLysLeuLysAspAsnLysGlnMet 515
     :::::::|||::::::      |||
1840 ATCGATGAAATCGACCATGACTATG 1864

seq_name: gb_pat1:AR105328

seq_documentation_block:
LOCUS       AR105328     1557 bp    DNA             PAT       14-FEB-2001
DEFINITION  Sequence 2 from patent US 6096539.
ACCESSION   AR105328
VERSION     AR105328.1  GI:12818925
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1557)
  AUTHORS   Gomes,B.Charles, Kasof,G.M. and Prosser,J.Caroline.
  TITLE     Protein activator of apoptosis
  JOURNAL   Patent: US 6096539-A 2 01-AUG-2000;
FEATURES             Location/Qualifiers
     source          1..1557
                     /organism="unknown"
BASE COUNT      395 a    422 c    448 g    291 t      1 others
ORIGIN


alignment_scores:
           Quality:  395.00             Length:    499
             Ratio:  1.452                Gaps:     21
 Percent Similarity:  54.509   Percent Identity: 28.056

alignment_block:
 US-09-445-223-1 x AR105328   ..

 Align seg 1/1  to: AR105328  from: 1  to: 1557

      11 ProThrIleProTyrHisLysLeuAlaAspLeuArgTyrLeuSerArgGl 27
         |||   :::      ::::::|||   :::   :::   :::::::::||
      40 CCCTTGGTGTCCATCGAGGAACTGGAGAACCAGGAGCTCGTCGGCAAAGG 89

      27 yAlaSerGlyThrValSerSerAlaArgHisAlaAspTrpArgValGlnV 44
         |:::   |||||||||      |||:::|||      |||      :::|
      90 CGGGTTCGGCACAGTGTTCCGGGCGCAACATAGGAAGTGGGGCTACGATG 139

      44 alAlaValLysHisLeuHisIleHisThrProLeuLeuAspSerGluArg 60
```

```
     |||||||||||   ::::::
 140 TGGCGGTCAAGATCGTAAAC...........................TCG 162

  61 LysAspValLeuArgGluAlaGluIleLeuHisLysAlaArgPheSerTy 77
     |||   :::   ||||||::::::   :::   :::         :::::
 163 AAGGCGATATCCAGGGAGGTCAAGGCCATGGCAAGTCTGGATAACGAATT 212

  77 rIlePheProIleLeuGlyIleCysAsnGluProGluPhe.......... 90
     :::::::   :::   |||:::   ::::::   ::::::
 213 CGTGCTGCGCCTAGAAGGGGTTATCGAGAAGGTGAACTGGGACCAAGATC 262

  91 .....LeuGlyIleValThrGluTyrMetProAsnGlySerLeuAsnGlu 105
             ::::::||||||::::::|||   ||||||||||||:::
 263 CCAAGCCGGCTCTGGTGACTAAATTCATGGAGAACGGCTCCTTGTCGGGG 312

 106 LeuLeuHisArgLysThrGluTyrProAspValAlaTrpProLeuArgPh 122
     ||||||:::   :::         |||         |||||||||
 313 CTGCTGCAGTCCCAGTGCCCTYGGCCC.........TGGCCGCTCCTTTG 353

 122 eArgIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisAsnMetT 139
      |||:::|||   |||::::::||||||:::   |||||||||:::::::
 354 CCGCCTGCTGAAAGAAGTGGTGCTTGGGATGTTTTACCTGCACGACCAGA 403

 139 hrProProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeuAsp 155
     ::|||   |||||||||:::|||||||||   :::|||:::|||||||||
 404 ACCCGGTGCTCCTGCACCGGGACCTCAAGCCATCCAACGTCCTGCTGGAC 453

 156 AsnGluPheHisValLysIleAlaAspPheGlyLeuSerLysTrpArgMe 172
        |||:::|||||||||:::||||||||||||||||||   ::::::
 454 CCAGAGCTGCACGTCAAGCTGGCAGATTTTGGCCTGTCCACATTTCAG.. 501

 172 tMetSerLeuSerGlnSerArgSerSerLysSerAlaProGluGlyGlyT 189
         :::   |||||||||   ::::::::::::   |||   |||||||
 502 ....GGAGGCTCACAGTCAGGGACAGGGTCCGGGGAGCCA...GGGGGCA 544

 189 hrIleIleTyrMetProProGluAsnTyrGluProGlyGlnLysSerArg 205
     ||:::   |||:::   ||||||   :::      :::   :::   :::
 545 CCCTGGGCTACTTGGCCCCAGAACTGTTT...GTTAACGTAAACCGGAAG 591

 206 AlaSerIleLysHisAspIleTyrSerTyrAlaValIleThrTrpGluVa 222
     ||||||         |||:::||||||::::::::::::   |||   ||
 592 GCCTCCACAGCCAGTGACGTCTACAGCTTCGGGATCCTAATGTGGGCAGT 641

 222 lLeuSerArgLysGlnProPheGluAspValThrAsnProLeuGlnIleM 239
     ||||:::   ::::::      |||      |||:::|||      :::
 642 GCTTGCTGGAAGAGAA...GTTGAGTTGCCAACCGAACCATCACTCGTGT 688

 239 etTyrSerVal...SerGlnGlyHisArgProValIleAsnGluGluSer 254
          :::|||   ::::::   :::||||||   :::   |||
 689 ACGAAGCAGTGTGCAACAGGCAGAACCGGCCTTCATTGGCTGAG...... 732

 255 LeuProTyr.........AspIleProHisArgAlaArgMetIleSerLe 268
     ||||||            :::   |||            :::   :::||
 733 CTGCCCCAAGCCGGGCCTGAGACTCCCGGCTTAGAAGGACTGAAGGAGCT 782

 268 uIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerPheLeuL 285
     |::::::      |||:::::::::|||   :::||||||||||||   :
 783 AATGCAGCTCTGCTGGAGCAGTGAGCCCAAGGACAGACCCTCCTTCCAGG 832

 285 ysCysLeuIleGluLeuGluProValLeuArgThrPheGluGluIleThr 301
     ::||||||   :::   :::   |||::::::      |||      :::
 833 AATGCCTACCAAAAACTGATGAAGTCTTCCAGATGGTGGAG......AAC 876

 302 PheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSerValSe 318
        ::::::|||:::::::::      |||      |||:::::::::
 877 AATATGAATGCTGCTGTCTCCACGGTAAAGGATTTCCTGTCTCAGCTCAG 926

 318 rSer.AlaIle......HisLeuCysAspLysLysLysMetGlu...... 330
      ||| ||||||         |||   :::::::::      |||
 927 GAGCAGCAATAGGAGATTTTCTATCCCAGAGTCAGGCCAAGGAGGGACAG 976

 331 ......LeuSerLeuAsnIlePro.....................ValAs 338
           ::::::|||:::   |||                     :::
 977 AAATGGATGGCTTTAGGAGAACCATAGAAAACCAGCACTCTCGTAATGAT 1026

 338 n.HisGly............ProGlnGluGluSerCysGlySerSerGln 350
       ||||||               |||   ||||||   |||:::||||||
1027 GTCATGGTTTCTGAGTGGCTAAACAAACTGAATCTAGAGGAGCCTCCCAG 1076

 351 LeuHisGluAsnSerGlySerProGluThrSerArgSerLeuProAlaPr 367
     |||                           |||   ::::::|||   ||
1077 CTC.....................TGTTCCTAAAAAATGCCCGAGCC 1102

 367 oGlnAspAsnAspPheLeuSerArgLysAla................... 377
     |   :::::::::      :::|||   |||
1103 TTACCAAGAGGAGCAGGGCACAAGAGGAGCAGGTTCCACAAGCCTGGACA 1152

 378 ............................................GlnAsp 379
                                                 :::|||
1153 GCAGGCACATCTTCAGATTCGATGGCCCAACCTCCCCAGACTCCAGAGAC 1202

 380 CysTyrPheMetLysLeuHisHisCysPro............GlyAsnHi 392
           |||:::|||            |||            ::::::
1203 CTCAACTTTCAGAAACCAGATGCCCAGCCCTACCTCAACTGGAACACCAA 1252

 392 sSerTrpAspSerThrIleSerGlySerGlnArgAlaAlaPheCysAspH 409
      ||||||            ||||||:::   :::
1253 GTCCTGGACCCCGAGGGAATCAGGGGGCTGAGAG................ 1286

 409 isLysThrThrProCysSerSerAlaIleIleAsnProLeuSerThrAla 425
             |||      :::   ::::::   :::|||   :::   :::
1287 ........ACAAGGCATGAACTGGTCCTGCAGGACCCCGGAGCCAAATCC 1328

 426 GlyAsnSerGluArgLeuGlnProGlyIleAlaGlnGlnTrpIleGlnSe 442
     :::|||            :::      |||         |||      ||
1329 AGTAACAGGGCGACCGCTCGTTAACATATACAACTGCTCTGGGGTGCAAG 1378

 442 rLysArgGluAspIleValAsnGlnMetThrGluAlaCysLeu 456
     |   |||:::::::::::::::      |||:::   ||||||
1379 TTGGAGACAACAACTACTTGACTATGCAACAGACAACTGCCTT 1421

seq_name: gb_pat2:AX067677

seq_documentation_block:
LOCUS       AX067677     1557 bp    DNA             PAT       19-JAN-2001
DEFINITION  Sequence 2 from Patent WO0077200.
ACCESSION   AX067677
VERSION     AX067677.1  GI:12329571
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1557)
  AUTHORS   Gomes,B.C., Kasof,G.M. and Prosser,J.C.
  TITLE     Receptor interacting protein rip3
  JOURNAL   Patent: WO 0077200-A 2 21-DEC-2000;
            AstraZeneca AB (SE)
FEATURES             Location/Qualifiers
     source          1..1557
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      395 a    422 c    448 g    291 t      1 others
ORIGIN


alignment_scores:
           Quality:  395.00             Length:    499
             Ratio:  1.452                Gaps:     21
 Percent Similarity:  54.509   Percent Identity: 28.056

alignment_block:
 US-09-445-223-1 x AX067677     ..

 Align seg 1/1  to: AX067677  from: 1  to: 1557
```

```
  11 ProThrIleProTyrHisLysLeuAlaAspLeuArgTyrLeuSerArgGl 27
     |||   :::      :::::::|||  :::   :::   :::::::::||
  40 CCCTTGGTGTCCATCGAGGAACTGGAGAACCAGGAGCTCGTCGGCAAAGG 89

  27 yAlaSerGlyThrValSerSerAlaArgHisAlaAspTrpArgValGlnV 44
     |:::   |||||||||      |||:::|||      |||      :::|
  90 CGGGTTCGGCACAGTGTTCCGGGCGCAACATAGGAAGTGGGGCTACGATG 139

  44 alAlaValLysHisLeuHisIleHisThrProLeuLeuAspSerGluArg 60
     |||||||||||   ::::::
 140 TGGCGGTCAAGATCGTAAAC...........................TCG 162

  61 LysAspValLeuArgGluAlaGluIleLeuHisLysAlaArgPheSerTy 77
     |||   :::   ||||||::::::   :::   :::         :::::
 163 AAGGCGATATCCAGGGAGGTCAAGGCCATGGCAAGTCTGGATAACGAATT 212

  77 rIlePheProIleLeuGlyIleCysAsnGluProGluPhe.......... 90
     :::::::   :::   |||:::   ::::::   ::::::
 213 CGTGCTGCGCCTAGAAGGGGTTATCGAGAAGGTGAACTGGGACCAAGATC 262

  91 .....LeuGlyIleValThrGluTyrMetProAsnGlySerLeuAsnGlu 105
             ::::::||||||::::::|||   ||||||||||||:::
 263 CCAAGCCGGCTCTGGTGACTAAATTCATGGAGAACGGCTCCTTGTCGGGG 312

 106 LeuLeuHisArgLysThrGluTyrProAspValAlaTrpProLeuArgPh 122
     ||||||:::   :::         |||         |||||||||
 313 CTGCTGCAGTCCCAGTGCCCTYGGCCC.........TGGCCGCTCCTTTG 353

 122 eArgIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisAsnMetT 139
      |||:::|||   |||::::::||||||:::   |||||||||:::::::
 354 CCGCCTGCTGAAAGAAGTGGTGCTTGGGATGTTTTACCTGCACGACCAGA 403

 139 hrProProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeuAsp 155
     ::|||   |||||||||:::|||||||||   :::|||:::|||||||||
 404 ACCCGGTGCTCCTGCACCGGGACCTCAAGCCATCCAACGTCCTGCTGGAC 453

 156 AsnGluPheHisValLysIleAlaAspPheGlyLeuSerLysTrpArgMe 172
        |||:::|||||||||:::||||||||||||||||||   ::::::
 454 CCAGAGCTGCACGTCAAGCTGGCAGATTTGGCCTGTCCACATTTCAG.. 501

 172 tMetSerLeuSerGlnSerArgSerSerLysSerAlaProGluGlyGlyT 189
         :::   |||||||||   ::::::::::::   |||   |||||||
 502 ....GGAGGCTCACAGTCAGGGACAGGGTCCGGGGAGCCA...GGGGGCA 544

 189 hrIleIleTyrMetProProGluAsnTyrGluProGlyGlnLysSerArg 205
     ||:::   |||:::   ||||||   :::      :::   :::   :::
 545 CCCTGGGCTACTTGGCCCCAGAACTGTTT...GTTAACGTAAACCGGAAG 591

 206 AlaSerIleLysHisAspIleTyrSerTyrAlaValIleThrTrpGluVa 222
     ||||||         |||:::||||||::::::::::::   |||   ||
 592 GCCTCCACAGCCAGTGACGTCTACAGCTTCGGGATCCTAATGTGGGCAGT 641

 222 lLeuSerArgLysGlnProPheGluAspValThrAsnProLeuGlnIleM 239
     ||||:::   ::::::      |||      |||:::|||      :::
 642 GCTTGCTGGAAGAGAA...GTTGAGTTGCCAACCGAACCATCACTCGTGT 688

 239 etTyrSerVal...SerGlnGlyHisArgProValIleAsnGluGluSer 254
          :::|||   ::::::   :::||||||   :::   |||
 689 ACGAAGCAGTGTGCAACAGGCAGAACCGGCCTTCATTGGCTGAG...... 732

 255 LeuProTyr.........AspIleProHisArgAlaArgMetIleSerLe 268
     ||||||            :::   |||            :::   :::||
 733 CTGCCCCAAGCCGGGCCTGAGACTCCCGGCTTAGAAGGACTGAAGGAGCT 782

 268 uIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerPheLeuL 285
     |::::::      |||:::::::::|||   :::||||||||||||   :
 783 AATGCAGCTCTGCTGGAGCAGTGAGCCCAAGGACAGACCCTCCTTCCAGG 832

 285 ysCysLeuIleGluLeuGluProValLeuArgThrPheGluGluIleThr 301
     ::||||||   :::   :::   |||::::::      |||      :::
 833 AATGCCTACCAAAAACTGATGAAGTCTTCCAGATGGTGGAG......AAC 876

 302 PheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSerValSe 318
        ::::::|||:::::::::      |||      |||:::::::::::
 877 AATATGAATGCTGCTGTCTCCACGGTAAAGGATTTCCTGTCTCAGCTCAG 926

 318 rSer.AlaIle......HisLeuCysAspLysLysLysMetGlu...... 330
      ||| ||||||         |||   :::::::::      |||
 927 GAGCAGCAATAGGAGATTTTCTATCCCAGAGTCAGGCCAAGGAGGGACAG 976

 331 ......LeuSerLeuAsnIlePro.....................ValAs 338
           ::::::|||:::   |||                     :::
 977 AAATGGATGGCTTTAGGAGAACCATAGAAAACCAGCACTCTCGTAATGAT 1026

 338 n.HisGly............ProGlnGluGluSerCysGlySerSerGln 350
       ||||||               |||   ||||||   |||:::||||||
1027 GTCATGGTTTCTGAGTGGCTAAACAAACTGAATCTAGAGGAGCCTCCCAG 1076

 351 LeuHisGluAsnSerGlySerProGluThrSerArgSerLeuProAlaPr 367
     |||                           |||   ::::::|||   ||
1077 CTC........................TGTTCCTAAAAAATGCCCGAGCC 1102

 367 oGlnAspAsnAspPheLeuSerArgLysAla................... 377
     |   :::::::::      :::|||   |||
1103 TTACCAAGAGGAGCAGGGCACAAGAGGAGCAGGTTCCACAAGCCTGGACA 1152

 378 ............................................GlnAsp 379
                                                 :::|||
1153 GCAGGCACATCTTCAGATTCGATGGCCCAACCTCCCCAGACTCCAGAGAC 1202

 380 CysTyrPheMetLysLeuHisHisCysPro............GlyAsnHi 392
           |||:::|||            |||            ::::::
1203 CTCAACTTTCAGAAACCAGATGCCCAGCCCTACCTCAACTGGAACACCAA 1252

 392 sSerTrpAspSerThrIleSerGlySerGlnArgAlaAlaPheCysAspH 409
      ||||||            ||||||:::   :::
1253 GTCCTGGACCCCGAGGAATCAGGGGGCTGAGAG................ 1286

 409 isLysThrThrProCysSerSerAlaIleIleAsnProLeuSerThrAla 425
             |||      :::   ::::::   :::|||   :::   :::
1287 ........ACAAGGCATGAACTGGTCCTGCAGGACCCCGGAGCCAAATCC 1328

 426 GlyAsnSerGluArgLeuGlnProGlyIleAlaGlnGlnTrpIleGlnSe 442
     :::|||            :::      |||         |||      ||
1329 AGTAACAGGGCGACCGCTCGTTAACATATACAACTGCTCTGGGGTGCAAG 1378

 442 rLysArgGluAspIleValAsnGlnMetThrGluAlaCysLeu 456
     |   |||:::::::::::::::      |||:::   ||||||
1379 TTGGAGACAACAACTACTTGACTATGCAACAGACAACTGCCTT 1421

seq_name: gb_pat1:AR105327

seq_documentation_block:
LOCUS       AR105327      1873 bp    DNA             PAT       14-FEB-2001
DEFINITION  Sequence 1 from patent US 6096539.
ACCESSION   AR105327
VERSION     AR105327.1  GI:12818924
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1873)
  AUTHORS   Gomes,B.Charles, Kasof,G.M. and Prosser,J.Caroline.
  TITLE     Protein activator of apoptosis
  JOURNAL   Patent: US 6096539-A 1 01-AUG-2000;
FEATURES             Location/Qualifiers
     source          1..1873
                     /organism="unknown"
BASE COUNT      471 a    531 c    518 g    353 t
ORIGIN


alignment_scores:
           Quality:  392.50             Length:    501
             Ratio:  1.443                Gaps:     22
 Percent Similarity:  54.291   Percent Identity: 28.144
```

```
alignment_block:
 US-09-445-223-1 x AR105327   ..

 Align seg 1/1  to: AR105327  from: 1  to: 1873

   11 ProThrIleProTyrHisLysLeuAlaAspLeuArgTyrLeuSerArgGl 27
      |||   :::      ::::::|||   :::   :::   :::::::::||
  204 CCCTTGGTGTCCATCGAGGAACTGGAGAACCAGGAGCTCGTCGGCAAAGG 253

   27 yAlaSerGlyThrValSerSerAlaArgHisAlaAspTrpArgValGlnV 44
      |:::   |||||||||      |||:::|||      |||      :::|
  254 CGGGTTCGGCACAGTGTTCCGGGCGCAACATAGGAAGTGGGGCTACGATG 303

   44 alAlaValLysHisLeuHisIleHisThrProLeuLeuAspSerGluArg 60
      |||||||||||   ::::::
  304 TGGCGGTCAAGATCGTAAAC...........................TCG 326

   61 LysAspValLeuArgGluAlaGluIleLeuHisLysAlaArgPheSerTy 77
      |||   :::   ||||||::::::   :::   :::         :::::
  327 AAGGCGATATCCAGGGAGGTCAAGGCCATGGCAAGTCTGGATAACGAATT 376

   77 rIlePheProIleLeuGlyIle.....................CysAsnG 87
      :::::::   :::   |||:::                        ::::
  377 CGTGCTGCGCCTAGAAGGGGTTATCGAGAAGGTCGGCGGCTCGAGCCAAG 426

   87 luProGluPheLeuGlyIleValThrGluTyrMetProAsnGlySerLeu 103
      ::|||:::      ::::::||||||::::::|||   ||||||||||||
  427 ATCCCAAG...CCGGCTCTGGTGACTAAATTCATGGAGAACGGCTCCTTG 473

  104 AsnGluLeuLeuHisArgLysThrGluTyrProAspValAlaTrpProLe 120
      :::   ||||||:::   :::         |||         ||||||||
  474 TCGGGGCTGCTGCAGTCCCAGTGCCCTCGGCCC.........TGGCCGCT 514

  120 uArgPheArgIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisA 137
      |      |||:::|||   |||::::::||||||:::   |||||||||:
  515 CCTTTGCCGCCTGCTGAAAGAAGTGGTGCTTGGGATGTTTTACCTGCACG 564

  137 snMetThrProProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeu 153
      ::::::::|||   |||||||||:::|||||||||   :::|||:::|||
  565 ACCAGAACCCGGTGCTCCTGCACCGGGACCTCAAGCCATCCAACGTCCTG 614

  154 LeuAspAsnGluPheHisValLysIleAlaAspPheGlyLeuSerLysTr 170
      ||||||   |||:::|||||||||:::||||||||||||||||||   ::
  615 CTGGACCCAGAGCTGCACGTCAAGCTGGCAGATTTTGGCCTGTCCACATT 664

  170 pArgMetMetSerLeuSerGlnSerArgSerSerLysSerAlaProGluG 187
      ::::      :::   |||||||||   ::::::::::::   |||   |
  665 TCAG......GGAGGCTCACAGTCAGGGACAGGGTCCGGGGAGCCA...G 705

  187 lyGlyThrIleIleTyrMetProProGluAsnTyrGluProGlyGlnLys 203
      ||||||||:::   |||:::   ||||||   :::      :::   :::
  706 GGGGCACCCTGGGCTACTTGGCCCCAGAACTGTTT...GTTAACGTAAAC 752

  204 SerArgAlaSerIleLysHisAspIleTyrSerTyrAlaValIleThrTr 220
         :::||||||         |||:::||||||::::::::::::   ||
  753 CGGAAGGCCTCCACAGCCAGTGACGTCTACAGCTTCGGGATCCTAATGTG 802

  220 pGluValLeuSerArgLysGlnProPheGluAspValThrAsnProLeuG 237
      |   ||||||:::   ::::::      |||      |||:::|||
  803 GGCAGTGCTTGCTGGAAGAGAA...GTTGAGTTGCCAACCGAACCATCAC 849

  237 lnIleMetTyrSerVal...SerGlnGlyHisArgProValIleAsnGlu 252
        :::      :::|||   ::::::   :::||||||   :::   |||
  850 TCGTGTACGAAGCAGTGTGCAACAGGCAGAACCGGCCTTCATTGGCTGAG 899

  253 GluSerLeuProTyr.........AspIleProHisArgAlaArgMetIl 266
            ||||||            :::   |||            :::
  900 ......CTGCCCCAAGCCGGGCCTGAGACTCCCGGCTTAGAAGGACTGAA 943

  266 eSerLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerP 283
       :::|||::::::      |||:::::::::|||   :::||||||||||
  944 GGAGCTAATGCAGCTCTGCTGGAGCAGTGAGCCCAAGGACAGACCCTCCT 993

  283 heLeuLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGlu 299
      ||   :::||||||   :::   :::   |||::::::      |||
  994 TCCAGGAATGCCTACCAAAAACTGATGAAGTCTTCCAGATGGTGGAG... 1040

  300 IleThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSe 316
         :::   ::::::|||:::::::::      |||      |||:::::
 1041 ...AACAATATGAATGCTGCTGTCTCCACGGTAAAGGATTTCCTGTCTCA 1087

  316 rValSerSer.AlaIle......HisLeuCysAspLysLysLysMetGlu 330
      ::::   ||| ||||||         |||   :::::::::      |||
 1088 GCTCAGGAGCAGCAATAGGAGATTTTCTATCCCAGAGTCAGGCCAAGGAG 1137

  331 ............LeuSerLeuAsnIlePro.................... 336
                  ::::::|||:::   |||
 1138 GGACAGAAATGGATGGCTTTAGGAGAACCATAGAAAACCAGCACTCTCGT 1187

  337 .ValAsn.HisGly............ProGlnGluGluSerCysGlySer 348
       :::    ||||||               |||   ||||||   |||:::
 1188 AATGATGTCATGGTTTCTGAGTGGCTAAACAAACTGAATCTAGAGGAGCC 1237

  349 SerGlnLeuHisGluAsnSerGlySerProGluThrSerArgSerLeuPr 365
      |||||||||                           |||   ::::::||
 1238 TCCCAGCTC........................TGTTCCTAAAAAATGCC 1263

  365 oAlaProGlnAspAsnAspPheLeuSerArgLysAla............. 377
      |   |||   :::::::::      :::|||   |||
 1264 CGAGCCTTACCAAGAGGAGCAGGGCACAAGAGGAGCAGGTTCCACAAGCC 1313

  377 .................................................. 377

 1314 TGGACAGCAGGCACATCTTCAGATTCGATGGCCCAACCTCCCCAGACTCC 1363

  378 GlnAspCysTyrPheMetLysLeuHisHisCysPro............Gl 390
      :::|||      |||:::|||            |||            ::
 1364 AGAGACCTCAACTTTCAGAAACCAGATGCCCAGCCCTACCTCAACTGGAA 1413

  390 yAsnHisSerTrpAspSerThrIleSerGlySerGlnArgAlaAlaPheC 407
      ::::   ||||||            ||||||:::   :::
 1414 CACCAAGTCCTGGACCCCGAGGAATCAGGGGGCTGAGAG........... 1453

  407 ysAspHisLysThrThrProCysSerSerAlaIleIleAsnProLeuSer 423
                    |||      :::   ::::::   :::|||   :::
 1454 ..............ACAAGGCATGAACTGGTCCTGCAGGACCCCGGAGCC 1489

  424 ThrAlaGlyAsnSerGluArgLeuGlnProGlyIleAlaGlnGlnTrpIl 440
         ::::::|||            :::      |||         |||
 1490 AAATCCAGTAACAGGGCGACCGCTCGTTAACATATACAACTGCTCTGGGG 1539

  440 eGlnSerLysArgGluAspIleValAsnGlnMetThrGluAlaCysLeu 456
          |||   |||:::::::::::::::      |||:::   ||||||
 1540 TGCAAGTTGGAGACAACAACTACTTGACTATGCAACAGACAACTGCCTT 1588

seq_name: gb_pat2:AX067676

seq_documentation_block:
LOCUS       AX067676     1873 bp    DNA             PAT       19-JAN-2001
DEFINITION  Sequence 1 from Patent WO0077200.
ACCESSION   AX067676
VERSION     AX067676.1  GI:12329570
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1873)
  AUTHORS   Gomes,B.C., Kasof,G.M. and Prosser,J.C.
  TITLE     Receptor interacting protein rip3
  JOURNAL   Patent: WO 0077200-A 1 21-DEC-2000;
            AstraZeneca AB (SE)
FEATURES             Location/Qualifiers
     source          1. .1873
                     /organism="Homo sapiens"
```

```
                    /db_xref="taxon:9606"
BASE COUNT      471 a    531 c    518 g    353 t
ORIGIN

alignment_scores:
           Quality:  392.50             Length:    501
             Ratio:  1.443                Gaps:     22
 Percent Similarity:  54.291   Percent Identity:  28.144

alignment_block:
 US-09-445-223-1 x AX067676    ..

 Align seg 1/1  to: AX067676  from: 1  to: 1873

  11 ProThrIleProTyrHisLysLeuAlaAspLeuArgTyrLeuSerArgGl 27
     |||   :::      ::::::|||   :::   :::   :::::::::||
 204 CCCTTGGTGTCCATCGAGGAACTGGAGAACCAGGAGCTCGTCGGCAAAGG 253

  27 yAlaSerGlyThrValSerSerAlaArgHisAlaAspTrpArgValGlnV 44
     |:::   |||||||||      |||:::|||      |||      :::|
 254 CGGGTTCGGCACAGTGTTCCGGGCGCAACATAGGAAGTGGGGCTACGATG 303

  44 alAlaValLysHisLeuHisIleHisThrProLeuLeuAspSerGluArg 60
     |||||||||||   ::::::
 304 TGGCGGTCAAGATCGTAAAC...........................TCG 326

  61 LysAspValLeuArgGluAlaGluIleLeuHisLysAlaArgPheSerTy 77
     |||   :::   ||||||::::::   :::   :::         :::::
 327 AAGGCGATATCCAGGGAGGTCAAGGCCATGGCAAGTCTGGATAACGAATT 376

  77 rIlePheProIleLeuGlyIle.....................CysAsnG 87
     :::::::   :::   |||:::                        ::::
 377 CGTGCTGCGCCTAGAAGGGGTTATCGAGAAGGTCGGCGGCTCGAGCCAAG 426

  87 luProGluPheLeuGlyIleValThrGluTyrMetProAsnGlySerLeu 103
     ::|||:::      ::::::||||||::::::|||   ||||||||||||
 427 ATCCCAAG...CCGGCTCTGGTGACTAAATTCATGGAGAACGGCTCCTTG 473

 104 AsnGluLeuLeuHisArgLysThrGluTyrProAspValAlaTrpProLe 120
     :::   ||||||:::   :::         |||         ||||||||
 474 TCGGGGCTGCTGCAGTCCCAGTGCCCTCGGCCC.........TGGCCGCT 514

 120 uArgPheArgIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisA 137
     |      |||:::|||   |||::::::||||||:::   |||||||||:
 515 CCTTTGCCGCCTGCTGAAAGAAGTGGTGCTTGGGATGTTTTACCTGCACG 564

 137 snMetThrProProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeu 153
     ::::::::|||   |||||||||:::|||||||||   :::|||:::|||
 565 ACCAGAACCCGGTGCTCCTGCACCGGGACCTCAAGCCATCCAACGTCCTG 614

 154 LeuAspAsnGluPheHisValLysIleAlaAspPheGlyLeuSerLysTr 170
     ||||||   |||:::|||||||||:::||||||||||||||||||   ::
 615 CTGGACCCAGAGCTGCACGTCAAGCTGGCAGATTTTGGCCTGTCCACATT 664

 170 pArgMetMetSerLeuSerGlnSerArgSerSerLysSerAlaProGluG 187
     ::::      :::   |||||||||   ::::::::::::   |||   |
 665 TCAG......GGAGGCTCACAGTCAGGGACAGGGTCCGGGGAGCCA...G 705

 187 lyGlyThrIleIleTyrMetProProGluAsnTyrGluProGlyGlnLys 203
     ||||||||:::   |||:::   ||||||   :::      :::   :::
 706 GGGGCACCCTGGGCTACTTGGCCCCAGAACTGTTT...GTTAACGTAAAC 752

 204 SerArgAlaSerIleLysHisAspIleTyrSerTyrAlaValIleThrTr 220
        :::||||||         |||:::||||||::::::::::::   ||
 753 CGGAAGGCCTCCACAGCCAGTGACGTCTACAGCTTCGGGATCCTAATGTG 802

 220 pGluValLeuSerArgLysGlnProPheGluAspValThrAsnProLeuG 237
     |   ||||||:::   ::::::      |||      |||:::|||
 803 GGCAGTGCTTGCTGGAAGAGAA...GTTGAGTTGCCAACCGAACCATCAC 849

 237 lnIleMetTyrSerVal...SerGlnGlyHisArgProValIleAsnGlu 252
       :::      :::|||   ::::::   :::||||||   :::   |||
 850 TCGTGTACGAAGCAGTGTGCAACAGGCAGAACCGGCCTTCATTGGCTGAG 899

 253 GluSerLeuProTyr.........AspIleProHisArgAlaArgMetIl 266
           ||||||            :::   |||            :::
 900 ......CTGCCCCAAGCCGGGCCTGAGACTCCCGGCTTAGAAGGACTGAA 943

 266 eSerLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerP 283
      :::|||::::::      |||:::::::::|||   :::||||||||||
 944 GGAGCTAATGCAGCTCTGCTGGAGCAGTGAGCCCAAGGACAGACCCTCCT 993

 283 heLeuLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGlu 299
     ||   :::||||||   :::   :::   |||::::::      |||
 994 TCCAGGAATGCCTACCAAAAACTGATGAAGTCTTCCAGATGGTGGAG... 1040

 300 IleThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSe 316
        :::   ::::::|||:::::::::      |||      |||:::::
1041 ...AACAATATGAATGCTGCTGTCTCCACGGTAAAGGATTTCCTGTCTCA 1087

 316 rValSerSer.AlaIle......HisLeuCysAspLysLysLysMetGlu 330
     ::::   ||| ||||||         |||   :::::::::      |||
1088 GCTCAGGAGCAGCAATAGGAGATTTTCTATCCCAGAGTCAGGCCAAGGAG 1137

 331 ............LeuSerLeuAsnIlePro.................... 336
                 ::::::|||:::   |||
1138 GGACAGAAATGGATGGCTTTAGGAGAACCATAGAAAACCAGCACTCTCGT 1187

 337 .ValAsn.HisGly............ProGlnGluGluSerCysGlySer 348
      :::    ||||||               |||   ||||||   |||:::
1188 AATGATGTCATGGTTTCTGAGTGGCTAAACAAACTGAATCTAGAGGAGCC 1237

 349 SerGlnLeuHisGluAsnSerGlySerProGluThrSerArgSerLeuPr 365
     |||||||||                           |||   ::::::||
1238 TCCCAGCTC...........................TGTTCCTAAAAAATGCC 1263

 365 oAlaProGlnAspAsnAspPheLeuSerArgLysAla............. 377
     |   |||   :::::::::      :::|||   |||
1264 CGAGCCTTACCAAGAGGAGCAGGGCACAAGAGGAGCAGGTTCCACAAGCC 1313

 377 .................................................. 377

1314 TGGACAGCAGGCACATCTTCAGATTCGATGGCCCAACCTCCCCAGACTCC 1363

 378 GlnAspCysTyrPheMetLysLeuHisHisCysPro............Gl 390
     :::|||      |||:::|||            |||            ::
1364 AGAGACCTCAACTTTCAGAAACCAGATGCCCAGCCCTACCTCAACTGGAA 1413

 390 yAsnHisSerTrpAspSerThrIleSerGlySerGlnArgAlaAlaPheC 407
     ::::   ||||||            ||||||:::   :::
1414 CACCAAGTCCTGGACCCCGAGGGAATCAGGGGGCTGAGAG.......... 1453

 407 ysAspHisLysThrThrProCysSerSerAlaIleIleAsnProLeuSer 423
                   |||      :::   ::::::   :::|||   :::
1454 ..............ACAAGGCATGAACTGGTCCTGCAGGACCCCGGAGCC 1489

 424 ThrAlaGlyAsnSerGluArgLeuGlnProGlyIleAlaGlnGlnTrpIl 440
        ::::::|||            :::      |||         |||
1490 AAATCCAGTAACAGGGCGACCGCTCGTTAACATATACAACTGCTCTGGGG 1539

 440 eGlnSerLysArgGluAspIleValAsnGlnMetThrGluAlaCysLeu 456
         |||   |||:::::::::::::::      |||:::   ||||||
1540 TGCAAGTTGGAGACAACAACTACTTGACTATGCAACAGACAACTGCCTT 1588
```

```
OM of: US-09-445-223-1 to: N_Geneseq_0401:*   out_format : pfs

Date: Jun 13, 2001  6:38 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn1_1/USPTO_spool/US09445223/runat_13062001_092033_29676/app_query.fasta_1.604
 -DB=N_Geneseq_0401 -QFMT=fastap -SUFFIX=p2n.rng -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -MINLEN=0 -MAXLEN=2000000000
 -USER=US09445223_@CGN1_1_0 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
 -WAIT -THREADS=1

Search information block:
Query: US-09-445-223-1
Query length: 540
Database: N_Geneseq_0401:*
Database sequences: 678276
Database length: 291890651
Search time (sec): 125.040000



score_list:
Sequence             Strd Orig    ZScore   EScore Len  ! Documentation   ..
/cgn1_8/gcgdata/geneseq/geneseqn/NA1999.DAT:X02558 + 2829.00 4448.48  7.6e-240  2098  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:Z46143 + 2823.00 4439.38  2.4e-239  2024  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:C77779 + 2823.00 4436.20  3.7e-239  2709  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA1999.DAT:Z09246 + 2817.00 4430.40  7.7e-239  1931  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:Z48762 + 2808.00 4413.34  6.9e-238  2502  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:F21841 +  797.00 1251.03  9.6e-62   463  ! H
/cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:Z61161 +  579.00  888.41  1.5e-41   2370  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:Z61784 +  579.00  884.11  2.6e-41   3516  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2001.DAT:C99717 +  579.00  884.11  2.6e-41   3516  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:Z61830 +  577.00  888.41  1.5e-41   1774  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2001.DAT:C99763 +  577.00  888.41  1.5e-41   1774  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:Z58584 +  577.00  885.60  2.2e-41   2294  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:Z61671 +  548.00  841.86  5.9e-39   1888  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2001.DAT:C99604 +  548.00  841.86  5.9e-39   1888  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA1997.DAT:T43752 +  397.00  601.03  1.5e-25   2268  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA1999.DAT:V99827 +  397.00  601.03  1.5e-25   2268  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:A47702 +  395.00  601.97  1.4e-25   1557  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:A47701 +  392.50  596.00  2.9e-25   1873  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA1999.DAT:V99818 +  391.00  586.05  1.0e-24   3750  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:A75675 +  388.00  587.43  8.8e-25   2140  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:Z50788 +  381.00  579.83  2.3e-24   1557  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA1994.DAT:Q57911 +  376.00  565.17  1.5e-23   2890  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA1995.DAT:Q98816 +  376.00  564.64  1.6e-23   3033  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA1997.DAT:T67126 +  376.00  564.64  1.6e-23   3033  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA1997.DAT:T43753 +  371.50  561.35  2.5e-23   2137  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA1997.DAT:T66408 +  370.00  559.61  3.1e-23   2016  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:Z89748 +  369.00  555.18  5.5e-23   2617  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA1998.DAT:V33547 +  367.00  579.54  2.4e-24   210  ! C
/cgn1_8/gcgdata/geneseq/geneseqn/NA2001.DAT:C90484 +  359.00  537.58  5.3e-22   3081  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2001.DAT:C90486 +  341.50  509.10  2.0e-20   3316  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:A47606 +  337.50  503.78  4.0e-20   3025  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:A47607 +  334.50  501.09  5.7e-20   2505  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2001.DAT:F44702 +  334.50  501.08  5.7e-20   2508  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2001.DAT:C90485 +  330.50  492.23  1.8e-19   3160  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:C48526 +  320.50  479.96  8.5e-19   2283  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:A79714 +  316.00  469.64  3.2e-18   3062  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:A47609 +  313.50  467.88  4.0e-18   2505  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:A47608 +  313.50  465.82  5.2e-18   3026  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA1999.DAT:X99698 +  313.00  471.29  2.6e-18   1704  !
/cgn1_8/gcgdata/geneseq/geneseqn/NA1999.DAT:X56285 +  312.00  469.18  3.4e-18   1788  !
```

```
/cgn1_8/gcgdata/geneseq/geneseqn/NA1997.DAT:T85095 +  312.00  468.19  3.9e-18   1959
/cgn1_8/gcgdata/geneseq/geneseqn/NA1999.DAT:X56279 +  312.00  464.87  5.9e-18   2656
/cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:A39105 +  312.00  464.87  5.9e-18   2656
/cgn1_8/gcgdata/geneseq/geneseqn/NA1999.DAT:X99696 +  312.00  464.35  6.3e-18   2785
/cgn1_8/gcgdata/geneseq/geneseqn/NA1991.DAT:Q10324 +  311.50  460.47  1.0e-17   3697

seq_name: /cgn1_8/gcgdata/geneseq/geneseqn/NA1999.DAT:X02558

seq_documentation_block:
ID   X02558 standard; cDNA; 2098 BP.
XX
AC   X02558;
XX
DT   07-MAY-1999  (first entry)
XX
DE   Human B1 cDNA.
XX
KW   B1 protein; intracellular mediator; modulator; inflammation; cell death;
KW   cell survival pathway; intracellular signalling; AIDS; cancer; human; ss.
XX
OS   Homo sapiens.
XX
PN   WO9855507-A2.
XX
PD   10-DEC-1998.
XX
PF   01-JUN-1998;   98WO-IL00255.
XX
PR   11-SEP-1997;   97IL-0121746.
PR   05-JUN-1997;   97IL-0121011.
PR   30-JUN-1997;   97IL-0121199.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Boldin M,  Malinin N,  Wallach D;
XX
DR   WPI; 1999-070258/06.
DR   P-PSDB; W92795.
XX
PT   New B1 protein regulates cell death and cell survival pathways -
PT   derivatives, DNA and antibodies, also regulate intracellular
PT   inflammation; for treating AIDS, cancer
XX
PS   Claim 4; Fig 3B; 90pp; English.
XX
CC   This invention describes the isolation of a novel human B1 protein which
CC   can interact with, intracellular mediators or modulators of inflammation,
CC   cell death and/or cell survival pathways, directly or indirectly. Cells
CC   can be modulated or mediated in inflammation, cell death or cell survival
CC   pathways or another intracellular signalling activity using B1.
CC   Conditions such as AIDS and cancer can be treated using B1. Antibodies,
CC   oligonucleotides and ribozymes can also be used to regulate the above
CC   pathways.
XX
SQ   Sequence 2098 BP; 649 A; 452 C; 449 G; 539 T; 9 other;


alignment_scores:
           Quality: 2829.00              Length:    540
             Ratio:  5.239                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-445-223-1 x X02558   ..

 Align seg 1/1  to: X02558  from: 1  to: 2098

       1 MetAsnGlyGluAlaIleCysSerAlaLeuProThrIleProTyrHisLy 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     260 ATGAACGGGGAGGCCATCTGCAGCGCCCTGCCCACCATTCCCTACCACAA 309

      17 sLeuAlaAspLeuArgTyrLeuSerArgGlyAlaSerGlyThrValSerS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
 310 ACTCGCCGACCTGCGCTACCTGAGCCGCGGCGCCTCTGGCACTGTGTCGT 359

  34 erAlaArgHisAlaAspTrpArgValGlnValAlaValLysHisLeuHis 50
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 360 CCGCCCGCCACGCAGACTGGCGCGTCCAGGTGGCCGTGAAGCACCTGCAC 409

  51 IleHisThrProLeuLeuAspSerGluArgLysAspValLeuArgGluAl 67
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 410 ATCCACACTCCGCTGCTCGACAGTGAAAGAAAGGATGTTTTAAGAGAAGC 459

  67 aGluIleLeuHisLysAlaArgPheSerTyrIlePheProIleLeuGlyI 84
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 460 TGAAATTTTACACAAAGCTAGATTTAGTTACATTTTTCCAATTTTGGGAA 509

  84 leCysAsnGluProGluPheLeuGlyIleValThrGluTyrMetProAsn 100
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 510 TTTGCAATGAGCCTGAATTTTTGGGAATAGTTACTGAATACATGCCAAAT 559

 101 GlySerLeuAsnGluLeuLeuHisArgLysThrGluTyrProAspValAl 117
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 560 GGATCATTAAATGAACTCCTACATAGGAAAACTGAATATCCTGATGTTGC 609

 117 aTrpProLeuArgPheArgIleLeuHisGluIleAlaLeuGlyValAsnT 134
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 610 TTGGCCATTGAGATTTCGCATCCTGCATGAAATTGCCCTTGGTGTAAATT 659

 134 yrLeuHisAsnMetThrProProLeuLeuHisHisAspLeuLysThrGln 150
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 660 ACCTGCACAATATGACTCCTCCTTTACTTCATCATGACTTGAAGACTCAG 709

 151 AsnIleLeuLeuAspAsnGluPheHisValLysIleAlaAspPheGlyLe 167
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 710 AATATCTTATTGGACAATGAATTTCATGTTAAGATTGCAGATTTTGGTTT 759

 167 uSerLysTrpArgMetMetSerLeuSerGlnSerArgSerSerLysSerA 184
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 760 ATCAAAGTGGCGCATGATGTCCCTCTCACAGTCACGAAGTAGCAAATCTG 809

 184 laProGluGlyGlyThrIleIleTyrMetProProGluAsnTyrGluPro 200
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 810 CACCAGAAGGAGGGACAATTATTTATATGCCACCTGAAAACTATGAACCT 859

 201 GlyGlnLysSerArgAlaSerIleLysHisAspIleTyrSerTyrAlaVa 217
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 860 GGACAAAAATCAAGGGCCAGTATCAAGCACGATATATATAGCTATGCAGT 909

 217 lIleThrTrpGluValLeuSerArgLysGlnProPheGluAspValThrA 234
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 910 TATCACATGGGAAGTGTTATCCAGAAAACAGCCTTTTGAAGATGTCACCA 959

 234 snProLeuGlnIleMetTyrSerValSerGlnGlyHisArgProValIle 250
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 960 ATCCTTTGCAGATAATGTATAGTGTGTCACAAGGACATCGACCTGTTATT 1009

 251 AsnGluGluSerLeuProTyrAspIleProHisArgAlaArgMetIleSe 267
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1010 AATGAAGAAAGTTTGCCATATGATATACCTCACCGAGCACGTATGATCTC 1059

 267 rLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerPheL 284
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1060 TCTAATAGAAAGTGGATGGGCACAAAATCCAGATGAAAGACCATCTTTCT 1109

 284 euLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGluIle 300
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1110 TAAAATGTTTAATAGAACTTGAACCAGTTTTGAGAACATTTGAAGAGATA 1159

 301 ThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSerVa 317
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1160 ACTTTTCTTGAAGCTGTTATTCAGCTAAAGAAAACAAAGTTACAGAGTGT 1209

 317 lSerSerAlaIleHisLeuCysAspLysLysLysMetGluLeuSerLeuA 334
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1210 TTCAAGTGCCATTCACCTATGTGACAAGAAGAAAATGGAATTATCTCTGA 1259
```

```
 334 snIleProValAsnHisGlyProGlnGluGluSerCysGlySerSerGln 350
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1260 ACATACCTGTAAATCATGGTCCACAAGAGGAATCATGTGGATCCTCTCAG 1309

 351 LeuHisGluAsnSerGlySerProGluThrSerArgSerLeuProAlaPr 367
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1310 CTCCATGAAAATAGTGGTTCTCCTGAAACTTCAAGGTCCCTGCCAGCTCC 1359

 367 oGlnAspAsnAspPheLeuSerArgLysAlaGlnAspCysTyrPheMetL 384
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1360 TCAAGACAATGATTTTTTATCTAGAAAAGCTCAAGACTGTTATTTTATGA 1409

 384 ysLeuHisHisCysProGlyAsnHisSerTrpAspSerThrIleSerGly 400
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1410 AGCTGCATCACTGTCCTGGAAATCACAGTTGGGATAGCACCATTTCTGGA 1459

 401 SerGlnArgAlaAlaPheCysAspHisLysThrThrProCysSerSerAl 417
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1460 TCTCAAAGGGCTGCATTCTGTGATCACAAGACCACTCCATGCTCTTCAGC 1509

 417 aIleIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeuGlnProG 434
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1510 AATAATAAATCCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAGCCTG 1559

 434 lyIleAlaGlnGlnTrpIleGlnSerLysArgGluAspIleValAsnGln 450
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1560 GTATAGCCCAGCAGTGGATCCAGAGCAAAAGGGAAGACATTGTGAACCAA 1609

 451 MetThrGluAlaCysLeuAsnGlnSerLeuAspAlaLeuLeuSerArgAs 467
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1610 ATGACAGAAGCCTGCCTTAACCAGTCGCTAGATGCCCTTCTGTCCAGGGA 1659

 467 pLeuIleMetLysGluAspTyrGluLeuValSerThrLysProThrArgT 484
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1660 CTTGATCATGAAAGAGGACTATGAACTTGTTAGTACCAAGCCTACAAGGA 1709

 484 hrSerLysValArgGlnLeuLeuAspThrThrAspIleGlnGlyGluGlu 500
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1710 CCTCAAAAGTCAGACAATTACTAGACACTACTGACATCCAAGGAGAAGAA 1759

 501 PheAlaLysValIleValGlnLysLeuLysAspAsnLysGlnMetGlyLe 517
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1760 TTTGCCAAAGTTATAGTACAAAAATTGAAAGATAACAAACAAATGGGTCT 1809

 517 uGlnProTyrProGluIleLeuValValSerArgSerProSerLeuAsnL 534
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1810 TCAGCCTTACCCGGAAATACTTGTGGTTTCTAGATCACCATCTTTAAATT 1859

 534 euLeuGlnAsnLysSerMet 540
     ||||||||||||||||||||
1860 TACTTCAAAATAAAAGCATG 1879

seq_name: /cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:Z46143

seq_documentation_block:
ID   Z46143 standard; cDNA; 2024 BP.
XX
AC   Z46143;
XX
DT   16-MAY-2000  (first entry)
XX
DE   cDNA sequence encoding a human phosphorylation effector PHSP-6.
XX
KW   Human; phosphorylation effector; PHSP; proliferative disorder;
KW   immune disorder; neuronal disorder; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             203..1825
FT                   /*tag=  a
FT                   /product= "phosphorylation effector"
```

```
XX
PN   WO200006728-A2.
XX
PD   10-FEB-2000.
XX
PF   28-JUL-1999;   99WO-US17132.
XX
PR   28-JUL-1998;   98US-0123494.
PR   14-SEP-1998;   98US-0152814.
PR   14-OCT-1998;   98US-0173482.
PR   03-NOV-1998;   98US-0106889.
PR   19-NOV-1998;   98US-0109093.
PR   22-DEC-1998;   98US-0113796.
PR   12-JAN-1999;   99US-0173482.
PR   12-JAN-1999;   99US-0229005.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Hillman JL,  Lal P,  Tang YT,  Corley NC,  Guegler KJ,  Baughn MR;
PI   Patterson C,  Bandman O,  Au-Young J,  Gorgone GA,  Yue H,  Azimzai Y;
PI   Reddy R,  Lu DAM,  Shih LL;
XX
DR   WPI; 2000-183125/16.
DR   P-PSDB; Y68774.
XX
PT   New human phosphorylation effectors useful for the diagnosis, treatment
PT   and prevention of proliferative, immune and neuronal disorders  -
XX
PS   Claim 9; Page 121-122; 142pp; English.
XX
CC   Z46138-Z46168 encode human phosphorylation effectors (PHSP),
CC   designated PHSP1-PHSP31 (the protein sequence for PHSP28 is not given
CC   in the specification). The sequences were isolated from cDNA libraries
CC   prepared from various human tissues. The PHSP proteins are useful for
CC   the diagnosis, treatment and prevention of proliferative disorders,
CC   immune disorders and neuronal disorders. The PHSP proteins form
CC   pharmaceutical compositions which useful for treating or preventing
CC   disorders associated with decreased PHSP expression/activity. PHSP
CC   antagonists are useful for treating or preventing disorders associated
CC   with increased PHSP expression/activity.
XX
SQ   Sequence 2024 BP; 612 A; 445 C; 434 G; 533 T; 0 other;


alignment_scores:
            Quality: 2823.00            Length:    540
              Ratio:   5.228              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.815

alignment_block:
 US-09-445-223-1 x Z46143   ..

 Align seg 1/1  to: Z46143  from: 1  to: 2024

       1 MetAsnGlyGluAlaIleCysSerAlaLeuProThrIleProTyrHisLy 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     203 ATGAACGGGGAGGCCATCTGCAGCGCCCTGCCCACCATTCCCTACCACAA 252

      17 sLeuAlaAspLeuArgTyrLeuSerArgGlyAlaSerGlyThrValSerS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     253 ACTCGCCGACCTGCGCTACCTGAGCCGCGGCGCCTCTGGCACTGTGTCGT 302

      34 erAlaArgHisAlaAspTrpArgValGlnValAlaValLysHisLeuHis 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     303 CCGCCCGCCACGCAGACTGGCGCGTCCAGGTGGCCGTGAAGCACCTGCAC 352

      51 IleHisThrProLeuLeuAspSerGluArgLysAspValLeuArgGluAl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     353 ATCCACACTCCGCTGCTCGACAGTGAAAGAAAGGATGTCTTAAGAGAAGC 402

      67 aGluIleLeuHisLysAlaArgPheSerTyrIlePheProIleLeuGlyI 84
         ||||||||||||||||||||||||||||||||||:::|||||||||||||
     403 TGAAATTTTACACAAAGCTAGATTTAGTTACATTCTTCCAATTTTGGGAA 452
```

```
      84 leCysAsnGluProGluPheLeuGlyIleValThrGluTyrMetProAsn 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     453 TTTGCAATGAGCCTGAATTTTTGGGAATAGTTACTGAATACATGCCAAAT 502

     101 GlySerLeuAsnGluLeuLeuHisArgLysThrGluTyrProAspValAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     503 GGATCATTAAATGAACTCCTACATAGGAAAACTGAATATCCTGATGTTGC 552

     117 aTrpProLeuArgPheArgIleLeuHisGluIleAlaLeuGlyValAsnT 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     553 TTGGCCATTGAGATTTCGCATCCTGCATGAAATTGCCCTTGGTGTAAATT 602

     134 yrLeuHisAsnMetThrProProLeuLeuHisHisAspLeuLysThrGln 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     603 ACCTGCACAATATGACTCCTCCTTTACTTCATCATGACTTGAAGACTCAG 652

     151 AsnIleLeuLeuAspAsnGluPheHisValLysIleAlaAspPheGlyLe 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     653 AATATCTTATTGGACAATGAATTTCATGTTAAGATTGCAGATTTTGGTTT 702

     167 uSerLysTrpArgMetMetSerLeuSerGlnSerArgSerSerLysSerA 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     703 ATCAAAGTGGCGCATGATGTCCCTCTCACAGTCACGAAGTAGCAAATCTG 752

     184 laProGluGlyGlyThrIleIleTyrMetProProGluAsnTyrGluPro 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     753 CACCAGAAGGAGGGACAATTATCTATATGCCACCTGAAAACTATGAACCT 802

     201 GlyGlnLysSerArgAlaSerIleLysHisAspIleTyrSerTyrAlaVa 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     803 GGACAAAAATCAAGGGCCAGTATCAAGCACGATATATATAGCTATGCAGT 852

     217 lIleThrTrpGluValLeuSerArgLysGlnProPheGluAspValThrA 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     853 TATCACATGGGAAGTGTTATCCAGAAAACAGCCTTTTGAAGATGTCACCA 902

     234 snProLeuGlnIleMetTyrSerValSerGlnGlyHisArgProValIle 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     903 ATCCTTTGCAGATAATGTATAGTGTGTCACAAGGACATCGACCTGTTATT 952

     251 AsnGluGluSerLeuProTyrAspIleProHisArgAlaArgMetIleSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     953 AATGAAGAAAGTTTGCCATATGATATACCTCACCGAGCACGTATGATCTC 1002

     267 rLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerPheL 284
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1003 TCTAATAGAAAGTGGATGGGCACAAAATCCAGATGAAAGACCATCTTTCT 1052

     284 euLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGluIle 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1053 TAAAATGTTTAATAGAACTTGAACCAGTTTTGAGAACATTTGAAGAGATA 1102

     301 ThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSerVa 317
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1103 ACTTTTCTTGAAGCTGTTATTCAGCTAAAGAAAACAAAGTTACAGAGTGT 1152

     317 lSerSerAlaIleHisLeuCysAspLysLysLysMetGluLeuSerLeuA 334
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1153 TTCAAGTGCCATTCACCTATGTGACAAGAAGAAAATGGAATTATCTCTGA 1202

     334 snIleProValAsnHisGlyProGlnGluGluSerCysGlySerSerGln 350
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1203 ACATACCTGTAAATCATGGTCCACAAGAGGAATCATGTGGATCCTCTCAG 1252

     351 LeuHisGluAsnSerGlySerProGluThrSerArgSerLeuProAlaPr 367
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1253 CTCCATGAAAATAGTGGTTCTCCTGAAACTTCAAGGTCCCTGCCAGCTCC 1302

     367 oGlnAspAsnAspPheLeuSerArgLysAlaGlnAspCysTyrPheMetL 384
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1303 TCAAGACAATGATTTTTTATCTAGAAAAGCTCAAGACTGTTATTTTATGA 1352
```

```
 384 ysLeuHisHisCysProGlyAsnHisSerTrpAspSerThrIleSerGly 400
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1353 AGCTGCATCACTGTCCTGGAAATCACAGTTGGGATAGCACCATTTCTGGA 1402

 401 SerGlnArgAlaAlaPheCysAspHisLysThrThrProCysSerSerAl 417
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1403 TCTCAAAGGGCTGCATTCTGTGATCACAAGACCACTCCATGCTCTTCAGC 1452

 417 aIleIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeuGlnProG 434
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1453 AATAATAAATCCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAGCCTG 1502

 434 lyIleAlaGlnGlnTrpIleGlnSerLysArgGluAspIleValAsnGln 450
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1503 GTATAGCCCAGCAGTGGATCCAGAGCAAAAGGGAAGACATTGTGAACCAA 1552

 451 MetThrGluAlaCysLeuAsnGlnSerLeuAspAlaLeuLeuSerArgAs 467
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1553 ATGACAGAAGCCTGCCTTAACCAGTCGCTAGATGCCCTTCTGTCCAGGGA 1602

 467 pLeuIleMetLysGluAspTyrGluLeuValSerThrLysProThrArgT 484
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1603 CTTGATCATGAAAGAGGACTATGAACTTGTTAGTACCAAGCCTACAAGGA 1652

 484 hrSerLysValArgGlnLeuLeuAspThrThrAspIleGlnGlyGluGlu 500
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1653 CCTCAAAAGTCAGACAATTACTAGACACTACTGACATCCAAGGAGAAGAA 1702

 501 PheAlaLysValIleValGlnLysLeuLysAspAsnLysGlnMetGlyLe 517
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1703 TTTGCCAAAGTTATAGTACAAAAATTGAAAGATAACAAACAAATGGGTCT 1752

 517 uGlnProTyrProGluIleLeuValValSerArgSerProSerLeuAsnL 534
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1753 TCAGCCTTACCCGGAAATACTTGTGGTTTCTAGATCACCATCTTTAAATT 1802

 534 euLeuGlnAsnLysSerMet 540
     ||||||||||||||||||||
1803 TACTTCAAAATAAAAGCATG 1822

seq_name: /cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:C77779

seq_documentation_block:
ID   C77779 standard; cDNA; 2709 BP.
XX
AC   C77779;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human cancer associated gene sequence SEQ ID NO:173.
XX
KW   Human; cancer associated gene; cancer antigen; detection; cancer;
KW   diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW   antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW   antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW   dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW   vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW   immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW   allergic reaction; graft versus host disease; organ rejection;
KW   haemostatic; thrombolytic; cardiovascular disorder; infection;
KW   neurological disease; drug screening; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055350-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05882.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
```

```
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587533/55.
DR   P-PSDB; B43570.
XX
PT   Novel isolated nucleic acids comprising sequences encoding peptides
PT   useful for treating or diagnosing e.g. cancer -
XX
PS   Claim 1; Page 751-752; 2352pp; English.
XX
CC   C77607 to C78448 encode the human cancer associated proteins given in
CC   B43398 to B44239. The proteins can have activities based on the tissues
CC   and cells the genes are expressed in. Example of activities include:
CC   cytostatic; proliferative; vulnerary; immunomodulator; antidiabetic;
CC   antiasthmatic; antirheumatic; antiarthritic; antiinflammatory;
CC   antithyroid; antiallergic; antibacterial; antiviral; dermatological;
CC   neuroprotective; cardiant; thrombolytic; coagulant; nootropic;
CC   vasotropic; antipsoriatic and antiangiogenic. The polynucleotides and
CC   polypeptides can be used for preventing, treating or ameliorating medical
CC   conditions and diagnosing pathological conditions. Polynucleotides,
CC   polypeptides, antibodies, agonists and antagonists from the present
CC   invention may be used to treat immune disorders by activating or
CC   inhibiting the proliferation, differentiation or mobilisation of immune
CC   cells, to treat disorders of haematopoietic cells, autoimmune disorders,
CC   allergic reactions, graft versus host disease and organ rejection,
CC   modulate haemostatic or thrombolytic activity, modulate inflammation,
CC   cancers, cardiovascular disorders, neurological disease and bacterial or
CC   viral infections. The peptides, nucleotides, antibodies, agonists and
CC   antagonists may be also be used in drug screens. C78449 to C78457 and
CC   B44240 represent sequences used in the exemplification of the present
CC   invention.
XX
SQ   Sequence 2709 BP; 810 A; 580 C; 540 G; 769 T; 10 other;


alignment_scores:
           Quality: 2823.00              Length:    540
             Ratio:  5.228                 Gaps:      0
 Percent Similarity: 100.000    Percent Identity: 99.815

alignment_block:
 US-09-445-223-1 x C77779   ..

 Align seg 1/1  to: C77779  from: 1  to: 2709

       1 MetAsnGlyGluAlaIleCysSerAlaLeuProThrIleProTyrHisLy 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     267 ATGAACGGGGAGGCCATCTGCAGCGCCCTGCCCACCATTCCCTACCACAA 316

      17 sLeuAlaAspLeuArgTyrLeuSerArgGlyAlaSerGlyThrValSerS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     317 ACTCGCCGACCTGCGCTACCTGAGCCGCGGCGCCTCTGGCACTGTGTCGT 366

      34 erAlaArgHisAlaAspTrpArgValGlnValAlaValLysHisLeuHis 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     367 CCGCCCGCCACGCAGACTGGCGCGTCCAGGTGGCCGTGAAGCACCTGCAC 416

      51 IleHisThrProLeuLeuAspSerGluArgLysAspValLeuArgGluAl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     417 ATCCACACTCCGCTGCTCGACAGTGAAAGAAAGGATGTCTTAAGAGAAGC 466

      67 aGluIleLeuHisLysAlaArgPheSerTyrIlePheProIleLeuGlyI 84
         ||||||||||||||||||||||||||||||||||:::|||||||||||||
     467 TGAAATTTACACAAAGCTAGATTTAGTTACATTCTTCCAATTTTGGGAA 516

      84 leCysAsnGluProGluPheLeuGlyIleValThrGluTyrMetProAsn 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     517 TTTGCAATGAGCCTGAATTTTTGGGAATAGTTACTGAATACATGCCAAAT 566

     101 GlySerLeuAsnGluLeuLeuHisArgLysThrGluTyrProAspValAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     567 GGATCATTAAATGAACTCCTACATAGGAAAACTGAATATCCTGATGTTGC 616
```

```
 117 aTrpProLeuArgPheArgIleLeuHisGluIleAlaLeuGlyValAsnT 134
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 617 TTGGCCATTGAGATTTCGCATCCTGCATGAAATTGCCCTTGGTGTAAATT 666

 134 yrLeuHisAsnMetThrProProLeuLeuHisHisAspLeuLysThrGln 150
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 667 ACCTGCACAATATGACTCCTCCTTTACTTCATCATGACTTGAAGACTCAG 716

 151 AsnIleLeuLeuAspAsnGluPheHisValLysIleAlaAspPheGlyLe 167
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 717 AATATCTTATTGGACAATGAATTTCATGTTAAGATTGCAGATTTTGGTTT 766

 167 uSerLysTrpArgMetMetSerLeuSerGlnSerArgSerSerLysSerA 184
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 767 ATCAAAGTGGCGCATGATGTCCCTCTCACAGTCACGAAGTAGCAAATCTG 816

 184 laProGluGlyGlyThrIleIleTyrMetProProGluAsnTyrGluPro 200
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 817 CACCAGAAGGAGGGACAATTATCTATATGCCACCTGAAAACTATGAACCT 866

 201 GlyGlnLysSerArgAlaSerIleLysHisAspIleTyrSerTyrAlaVa 217
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 867 GGACAAAAATCAAGGGCCAGTATCAAGCACGATATATATAGCTATGCAGT 916

 217 lIleThrTrpGluValLeuSerArgLysGlnProPheGluAspValThrA 234
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 917 TATCACATGGGAAGTGTTATCCAGAAAACAGCCTTTTGAAGATGTCACCA 966

 234 snProLeuGlnIleMetTyrSerValSerGlnGlyHisArgProValIle 250
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 967 ATCCTTTGCAGATAATGTATAGTGTGTCACAAGGACATCGACCTGTTATT 1016

 251 AsnGluGluSerLeuProTyrAspIleProHisArgAlaArgMetIleSe 267
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1017 AATGAAGAAAGTTTGCCATATGATATACCTCACCGAGCACGTATGATCTC 1066

 267 rLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerPheL 284
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1067 TCTAATAGAAAGTGGATGGGCACAAAATCCAGATGAAAGACCATCTTTCT 1116

 284 euLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGluIle 300
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1117 TAAAATGTTTAATAGAACTTGAACCAGTTTTGAGAACATTTGAAGAGATA 1166

 301 ThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSerVa 317
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1167 ACTTTTCTTGAAGCTGTTATTCAGCTAAAGAAAACAAAGTTACAGAGTGT 1216

 317 lSerSerAlaIleHisLeuCysAspLysLysLysMetGluLeuSerLeuA 334
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1217 TTCAAGTGCCATTCACCTATGTGACAAGAAGAAAATGGAATTATCTCTGA 1266

 334 snIleProValAsnHisGlyProGlnGluGluSerCysGlySerSerGln 350
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1267 ACATACCTGTAAATCATGGTCCACAAGAGGAATCATGTGGATCCTCTCAG 1316

 351 LeuHisGluAsnSerGlySerProGluThrSerArgSerLeuProAlaPr 367
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1317 CTCCATGAAAATAGTGGTTCTCCTGAAACTTCAAGGTCCCTGCCAGCTCC 1366

 367 oGlnAspAsnAspPheLeuSerArgLysAlaGlnAspCysTyrPheMetL 384
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1367 TCAAGACAATGATTTTTTATCTAGAAAAGCTCAAGACTGTTATTTTATGA 1416

 384 ysLeuHisHisCysProGlyAsnHisSerTrpAspSerThrIleSerGly 400
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1417 AGCTGCATCACTGTCCTGGAAATCACAGTTGGGAYAGCACCATTTCTGGA 1466

 401 SerGlnArgAlaAlaPheCysAspHisLysThrThrProCysSerSerAl 417
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1467 TCTCAAAGGGCTGCATTCTGTGATCACAAGACCACTCCATGCTCTTCAGC 1516
```

```
 417 aIleIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeuGlnProG 434
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1517 AATAATAAATCCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAGCCTG 1566

 434 lyIleAlaGlnGlnTrpIleGlnSerLysArgGluAspIleValAsnGln 450
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1567 GTATAGCCCAGCAGTGGATCCAGAGCAAAAGGGAAGACATTGTGAACCAA 1616

 451 MetThrGluAlaCysLeuAsnGlnSerLeuAspAlaLeuLeuSerArgAs 467
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1617 ATGACAGAAGCCTGCCTTAACCAGTCGCTAGATGCCCTTCTGTCCAGGGA 1666

 467 pLeuIleMetLysGluAspTyrGluLeuValSerThrLysProThrArgT 484
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1667 CTTGATCATGAAAGAGGACTATGAACTTGTTAGTACCAAGCCTACAAGGA 1716

 484 hrSerLysValArgGlnLeuLeuAspThrThrAspIleGlnGlyGluGlu 500
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1717 CCTCAAAAGTCAGACAATTACTAGACACTACTGACATCCAAGGAGAAGAA 1766

 501 PheAlaLysValIleValGlnLysLeuLysAspAsnLysGlnMetGlyLe 517
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1767 TTTGCCAAAGTTATAGTACAAAAATTGAAAGATAACAAACAAATGGGTCT 1816

 517 uGlnProTyrProGluIleLeuValValSerArgSerProSerLeuAsnL 534
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1817 TCAGCCTTACCCGGAAATACTTGTGGTTTCTAGATCACCATCTTTAAATT 1866

 534 euLeuGlnAsnLysSerMet 540
     ||||||||||||||||||||
1867 TACTTCAAAATAAAAGCATG 1886
```

```
seq_name: /cgn1_8/gcgdata/geneseq/geneseqn/NA1999.DAT:Z09246

seq_documentation_block:
ID   Z09246 standard; cDNA; 1931 BP.
XX
AC   Z09246;
XX
DT   25-OCT-1999  (first entry)
XX
DE   Human CARD-3 cDNA.
XX
KW   CARD-3; caspase recruitment domain; CARD-4; regulation; detection;
KW   caspase activation; detection; screening; therapy; diagnosis; disease;
KW   apoptotic cell death; Fas/APO-1 receptor complex; TNF receptor complex;
KW   cancer; follicular lymphoma; carcinoma; p53 mutation; viral infection;
KW   hormone-dependent tumour; autoimmune disorder; Alzheimer's disease;
KW   systemic lupus erythematosis; immune-mediated glomerulonephritis; stroke;
KW   Parkinson's disease; amyotrophic lateral sclerosis; retinitis pigmentosa;
KW   spinal muscular dystrophy; cerebellar degeneration; anaemia; drug;
KW   myelodysplastic syndrome; myocardial infarction; cell proliferation;
KW   cell differentiation; cell survival; CARD-4L; CARD-4S; CARD-4Y;
KW   CARD-4Z; human; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             214..1836
FT                   /*tag=  a
FT                   /product= "CARD-3"
XX
PN   WO9940102-A1.
XX
PD   12-AUG-1999.
XX
PF   05-FEB-1999;   99WO-US02544.
XX
PR   08-DEC-1998;   98US-0207359.
PR   06-FEB-1998;   98US-0019942.
PR   17-JUN-1998;   98US-0099041.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
```

```
XX
PI   Bertin J;
XX
DR   WPI; 1999-494269/41.
DR   P-PSDB; Y31140.
XX
PT   Novel CARD-3 and CARD-4 genes and polypeptides used or treating
PT   regulation of cellular proliferation and differentiation and cell
PT   survival
XX
PS   Example 2; Fig 1; 181pp; English.
XX
CC   This invention describes the isolation of novel human caspase
CC   recruitment domain, CARD-3 and CARD-4 polynucleotides and proteins and a
CC   partial murine CARD-4L protein and genes. The genes and proteins of
CC   the invention are involved in the regulation of caspase activation.
CC   The caspase recruitment domain (CARD) polynucleotides, polypeptides,
CC   homologues and antibodies can be used in screening assays, detection
CC   assays, predictive medicine and therapeutic and prophylactic methods of
CC   treatment. The methods may be used to diagnose and treat patients which
CC   are suffering from a disorder associated with abnormal level or rate of
CC   apoptotic cell death, abnormal activity of the Fas/APO-1 receptor
CC   complex, abnormal activity of the TNF receptor complex, or abnormal
CC   activity of a caspase. Diseases that may be treated include cancer
CC   (particularly follicular lymphoma, carcinomas associated with mutations
CC   in p53 and hormone-dependent tumours), autoimmune disorders (e.g.
CC   systemic lupus erythematosis, immune-mediated glomerulonephritis), viral
CC   infections, Alzheimer's disease, Parkinson's disease, amyotrophic lateral
CC   sclerosis, retinitis pigmentosa, spinal muscular dystrophy, cerebellar
CC   degeneration, anaemia, myelodysplastic syndrome, myocardial infarction,
CC   and stroke. CARD-3 protein interacts with other cellular proteins, and so
CC   can be used for regulation of cellular proliferation and differentiation
CC   and cell survival. The CARD proteins may also be used to for screen drugs
CC   or compounds which modulate their activity. The CARD-4 gene can express a
CC   long transcript that encodes CARD-4L, a short transcript that encodes
CC   CARD-4S or two CARD-4 splice variants, CARD-4Y and CARD-4Z. This sequence
CC   encodes the human CARD-3 protein described in the method of the
CC   invention.
XX
SQ   Sequence 1931 BP; 613 A; 429 C; 416 G; 473 T; 0 other;


alignment_scores:
           Quality: 2817.00          Length:    540
             Ratio:  5.226             Gaps:      0
 Percent Similarity:  99.815   Percent Identity: 99.630

alignment_block:
 US-09-445-223-1 x Z09246   ..

 Align seg 1/1  to: Z09246  from: 1  to: 1931

       1 MetAsnGlyGluAlaIleCysSerAlaLeuProThrIleProTyrHisLy 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     214 ATGAACGGGGAGGCCATCTGCAGCGCCCTGCCCACCATTCCCTACCACAA 263

      17 sLeuAlaAspLeuArgTyrLeuSerArgGlyAlaSerGlyThrValSerS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     264 ACTCGCCGACCTGCGCTACCTGAGCCGCGGCGCCTCTGGCACTGTGTCGT 313

      34 erAlaArgHisAlaAspTrpArgValGlnValAlaValLysHisLeuHis 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     314 CCGCCCGCCACGCAGACTGGCGCGTCCAGGTGGCCGTGAAGCACCTGCAC 363

      51 IleHisThrProLeuLeuAspSerGluArgLysAspValLeuArgGluAl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     364 ATCCACACTCCGCTGCTCGACAGTGAAAGAAAGGATGTCTTAAGAGAAGC 413

      67 aGluIleLeuHisLysAlaArgPheSerTyrIlePheProIleLeuGlyI 84
         ||||||||||||||||||||||||||||||||||:::|||||||||||||
     414 TGAAATTTTACACAAAGCTAGATTTAGTTACATTCTTCCAATTTTGGGAA 463

      84 leCysAsnGluProGluPheLeuGlyIleValThrGluTyrMetProAsn 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     464 TTTGCAATGAGCCTGAATTTTTGGGAATAGTTACTGAATACATGCCAAAT 513

     101 GlySerLeuAsnGluLeuLeuHisArgLysThrGluTyrProAspValAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     514 GGATCATTAAATGAACTCCTACATAGGAAAACTGAATATCCTGATGTTGC 563

     117 aTrpProLeuArgPheArgIleLeuHisGluIleAlaLeuGlyValAsnT 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     564 TTGGCCATTGAGATTTCGTATCCTGCATGAAATTGCCCTTGGTGTAAATT 613

     134 yrLeuHisAsnMetThrProProLeuLeuHisHisAspLeuLysThrGln 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     614 ACCTGCACAATATGACTCCTCCTTTACTTCATCATGACTTGAAGACTCAG 663

     151 AsnIleLeuLeuAspAsnGluPheHisValLysIleAlaAspPheGlyLe 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     664 AATATCTTATTGGACAATGAATTTCATGTTAAGATTGCAGATTTTGGTTT 713

     167 uSerLysTrpArgMetMetSerLeuSerGlnSerArgSerSerLysSerA 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     714 ATCAAAGTGGCGCATGATGTCCCTCTCACAGTCACGAAGTAGCAAATCTG 763

     184 laProGluGlyGlyThrIleIleTyrMetProProGluAsnTyrGluPro 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     764 CACCAGAAGGAGGGACAATTATCTATATGCCACCTGAAAACTATGAACCT 813

     201 GlyGlnLysSerArgAlaSerIleLysHisAspIleTyrSerTyrAlaVa 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     814 GGACAAAAATCAAGGGCCAGTATCAAGCACGATATATATAGCTATGCAGT 863

     217 lIleThrTrpGluValLeuSerArgLysGlnProPheGluAspValThrA 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     864 TATCACATGGGAAGTGTTATCCAGAAAACAGCCTTTTGAAGATGTCACCA 913

     234 snProLeuGlnIleMetTyrSerValSerGlnGlyHisArgProValIle 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     914 ATCCTTTGCAGATAATGTATAGTGTGTCACAAGGACATCGACCTGTTATT 963

     251 AsnGluGluSerLeuProTyrAspIleProHisArgAlaArgMetIleSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     964 AATGAAGAAAGTTTGCCATATGATATACCTCACCGAGCACGTATGATCTC 1013

     267 rLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerPheL 284
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1014 TCTAATAGAAAGTGGATGGGCACAAAATCCAGATGAAAGACCATCTTTCT 1063

     284 euLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGluIle 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1064 TAAAATGTTTAATAGAACTTGAACCAGTTTTGAGAACATTTGAAGAGATA 1113

     301 ThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSerVa 317
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1114 ACTTTTCTTGAAGCTGTTATTCAGCTAAAGAAAACAAAGTTACAGAGTGT 1163

     317 lSerSerAlaIleHisLeuCysAspLysLysLysMetGluLeuSerLeuA 334
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1164 TTCAAGTGCCATTCACCTATGTGACAAGAAGAAAATGGAATTATCTCTGA 1213

     334 snIleProValAsnHisGlyProGlnGluGluSerCysGlySerSerGln 350
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1214 ACATACCTGTAAATCATGGTCCACAAGAGGAATCATGTGGATCCTCTCAG 1263

     351 LeuHisGluAsnSerGlySerProGluThrSerArgSerLeuProAlaPr 367
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1264 CTCCATGAAAATAGTGGTTCTCCTGAAACTTCAAGGTCCCTGCCAGCTCC 1313

     367 oGlnAspAsnAspPheLeuSerArgLysAlaGlnAspCysTyrPheMetL 384
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1314 TCAAGACAATGATTTTTTATCTAGAAAAGCTCAAGACTGTTATTTTATGA 1363

     384 ysLeuHisHisCysProGlyAsnHisSerTrpAspSerThrIleSerGly 400
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
  1364 AGCTGCATCACTGTCCTGGAAATCACAGTTGGGATAGCACCATTTCTGGA 1413

   401 SerGlnArgAlaAlaPheCysAspHisLysThrThrProCysSerSerAl 417
       |||||||||||||||||||||||||||||||||   ||||||||||||||
  1414 TCTCAAAGGGCTGCATTCTGTGATCACAAGACCATTCCATGCTCTTCAGC 1463

   417 aIleIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeuGlnProG 434
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1464 AATAATAAATCCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAGCCTG 1513

   434 lyIleAlaGlnGlnTrpIleGlnSerLysArgGluAspIleValAsnGln 450
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1514 GTATAGCCCAGCAGTGGATCCAGAGCAAAAGGGAAGACATTGTGAACCAA 1563

   451 MetThrGluAlaCysLeuAsnGlnSerLeuAspAlaLeuLeuSerArgAs 467
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1564 ATGACAGAAGCCTGCCTTAACCAGTCGCTAGATGCCCTTCTGTCCAGGGA 1613

   467 pLeuIleMetLysGluAspTyrGluLeuValSerThrLysProThrArgT 484
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1614 CTTGATCATGAAAGAGGACTATGAACTTGTTAGTACCAAGCCTACAAGGA 1663

   484 hrSerLysValArgGlnLeuLeuAspThrThrAspIleGlnGlyGluGlu 500
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1664 CCTCAAAAGTCAGACAATTACTAGACACTACTGACATCCAAGGAGAAGAA 1713

   501 PheAlaLysValIleValGlnLysLeuLysAspAsnLysGlnMetGlyLe 517
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1714 TTTGCCAAAGTTATAGTACAAAAATTGAAAGATAACAAACAAATGGGTCT 1763

   517 uGlnProTyrProGluIleLeuValValSerArgSerProSerLeuAsnL 534
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1764 TCAGCCTTACCCGGAAATACTTGTGGTTTCTAGATCACCATCTTTAAATT 1813

   534 euLeuGlnAsnLysSerMet 540
       ||||||||||||||||||||
  1814 TACTTCAAAATAAAAGCATG 1833

seq_name: /cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:Z48762

seq_documentation_block:
ID   Z48762 standard; cDNA; 2502 BP.
XX
AC   Z48762;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Human RICK coding sequence.
XX
KW   RICK; human; RIP-like interacting CLARP kinase; apoptosis regulator; ARC;
KW   caspase-8; caspase-10; CD95 signalling; apoptosis signalling pathway;
KW   CIDE-A; CIDE-B; DREP-1; diagnosis; cell growth; apoptosis dysregulation;
KW   gene therapy; cardiac disorder; cancer; neurodegenerative disorder; AIDS;
KW   aplastic anaemia; ischaemic injury; toxin-induced liver disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO9955134-A2.
XX
PD   04-NOV-1999.
XX
PF   27-APR-1999;   99WO-US09183.
XX
PR   27-APR-1998;   98US-0069023.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Nunez G,  Inohara N,  Koseki T;
XX
DR   WPI; 2000-072163/06.
DR   P-PSDB; Y59404.
XX
PT   Compositions for identifying apoptosis signalling pathway inhibitors
```

```
PT   useful for treating diseases -
XX
PS   Claim 8; Fig 7b; 93pp; English.
XX
CC   This sequence encodes the human RICK (RIP-like interacting CLARP kinase)
CC   protein of the invention. The RICK protein acts as a positive regulator
CC   of apoptosis, potentiating apoptosis induced by caspase-8 and caspase-10
CC   during CD95 signalling. The invention provides methods for identifying
CC   apoptosis signalling pathway inhibitors and activators, and methods and
CC   compositions for screening compounds which will modulate the interactions
CC   of the various compositions identified: ARC, RICK, and the CIDE family of
CC   activators (CIDE-A, CIDE-B and DREP-1). RICK is useful in screening
CC   assays for agents, useful in the diagnosis, prognosis or treatment of
CC   disease associated with excess cell growth and dysregulation of
CC   apoptosis. Complexes containing RICK and CLARP can be used in drug
CC   screening assays to identify inhibitor molecules blocking CD95-mediated
CC   apoptosis. Overexpression of ARC in an in vitro cell system can be used
CC   to identify inhibitors of the enzymatic activity of caspase-8.
CC   Identification of ARC-like inhibitory compounds may be useful for gene
CC   therapy treatment of disease with increased cell death in muscle tissue
CC   and cardiac disorders. Therapeutic compositions of CIDEs can be used to
CC   treat e.g. cancer, AIDS, neurodegenerative disorders, aplastic anaemia,
CC   ischaemic injury, and toxin-induced liver disease. AntiRICK antibodies
CC   can be used as reagents for the preparation or affinity chromatography
CC   media, and for diagnostically measuring RICK levels. A specific inhibitor
CC   of an essential step in the biochemistry of apoptosis is needed. RICK
CC   interaction with intracellular factors such as CLARP and FADD appears to
CC   be essential for apoptosis, inhibitors of RICK binding to intracellular
CC   apoptosis factors are potential drug candidates.
XX
SQ   Sequence 2502 BP; 769 A; 535 C; 499 G; 699 T; 0 other;


alignment_scores:
           Quality: 2808.00             Length:    541
             Ratio:  5.200                Gaps:      1
 Percent Similarity:  99.815   Percent Identity: 99.445

alignment_block:
 US-09-445-223-1 x Z48762   ..

 Align seg 1/1  to: Z48762  from: 1  to: 2502

       1 MetAsnGlyGluAlaIleCysSerAla.LeuProThrIleProTyrHisL 17
         ||||||||||||||||||||||||||| :::|||||||||||||||||||
     225 ATGAACGGGGAGGCCATCTGCAGCGCCCATGCCCACCATTCCCTACCACA 274

      17 ysLeuAlaAspLeuArgTyrLeuSerArgGlyAlaSerGlyThrValSer 33
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     275 AACTCGCCGACCTGCGCTACCTGAGCCGCGGCGCCTCTGGCACTGTGTCG 324

      34 SerAlaArgHisAlaAspTrpArgValGlnValAlaValLysHisLeuHi 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     325 TCCGCCCGCCACGCAGACTGGCGCGTCCAGGTGGCCGTGAAGCACCTGCA 374

      50 sIleHisThrProLeuLeuAspSerGluArgLysAspValLeuArgGluA 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     375 CATCCACACTCCGCTGCTCGACAGTGAAAGAAAGGATGTCTTAAGAGAAG 424

      67 laGluIleLeuHisLysAlaArgPheSerTyrIlePheProIleLeuGly 83
         |||||||||||||||||||||||||||||||||||:::||||||||||||
     425 CTGAAATTTTACACAAAGCTAGATTTAGTTACATTCTTCCAATTTTGGGA 474

      84 IleCysAsnGluProGluPheLeuGlyIleValThrGluTyrMetProAs 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     475 ATTTGCAATGAGCCTGAATTTTTGGGAATAGTTACTGAATACATGCCAAA 524

     100 nGlySerLeuAsnGluLeuLeuHisArgLysThrGluTyrProAspValA 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     525 TGGATCATTAAATGAACTCCTACATAGGAAAACTGAATATCCTGATGTTG 574

     117 laTrpProLeuArgPheArgIleLeuHisGluIleAlaLeuGlyValAsn 133
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
 575 CTTGGCCATTGAGATTTCGCATCCTGCATGAAATTGCCCTTGGTGTAAAT 624

 134 TyrLeuHisAsnMetThrProProLeuLeuHisHisAspLeuLysThrGl 150
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 625 TACCTGCACAATATGACTCCTCCTTTACTTCATCATGACTTGAAGACTCA 674

 150 nAsnIleLeuLeuAspAsnGluPheHisValLysIleAlaAspPheGlyL 167
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 675 GAATATCTTATTGGACAATGAATTTCATGTTAAGATTGCAGATTTTGGTT 724

 167 euSerLysTrpArgMetMetSerLeuSerGlnSerArgSerSerLysSer 183
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 725 TATCAAAGTGGCGCATGATGTCCCTCTCACAGTCACGAAGTAGCAAATCT 774

 184 AlaProGluGlyGlyThrIleIleTyrMetProProGluAsnTyrGluPr 200
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 775 GCACCAGAAGGAGGGACAATTATCTATATGCCACCTGAAAACTATGAACC 824

 200 oGlyGlnLysSerArgAlaSerIleLysHisAspIleTyrSerTyrAlaV 217
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 825 TGGACAAAAATCAAGGGCCAGTATCAAGCACGATATATATAGCTATGCAG 874

 217 alIleThrTrpGluValLeuSerArgLysGlnProPheGluAspValThr 233
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 875 TTATCACATGGGAAGTGTTATCCAGAAAACAGCCTTTTGAAGATGTCACC 924

 234 AsnProLeuGlnIleMetTyrSerValSerGlnGlyHisArgProValIl 250
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 925 AATCCTTTGCAGATAATGTATAGTGTGTCACAAGGACATCGACCTGTTAT 974

 250 eAsnGluGluSerLeuProTyrAspIleProHisArgAlaArgMetIleS 267
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 975 TAATGAAGAAAGTTTGCCATATGATATACCTCACCGAGCACGTATGATCT 1024

 267 erLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerPhe 283
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1025 CTCTAATAGAAAGTGGATGGGCACAAAATCCAGATGAAAGACCATCTTTC 1074

 284 LeuLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGluIl 300
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1075 TTAAAATGTTTAATAGAACTTGAACCAGTTTTGAGAACATTTGAAGAGAT 1124

 300 eThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSerV 317
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1125 AACTTTTCTTGAAGCTGTTATTCAGCTAAAGAAAACAAAGTTACAGAGTG 1174

 317 alSerSerAlaIleHisLeuCysAspLysLysLysMetGluLeuSerLeu 333
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1175 TTTCAAGTGCCATTCACCTATGTGACAAGAAGAAAATGGAATTATCTCTG 1224

 334 AsnIleProValAsnHisGlyProGlnGluGluSerCysGlySerSerGl 350
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1225 AACATACCTGTAAATCATGGTCCACAAGAGGAATCATGTGGATCCTCTCA 1274

 350 nLeuHisGluAsnSerGlySerProGluThrSerArgSerLeuProAlaP 367
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1275 GCTCCATGAAAATAGTGGTTCTCCTGAAACTTCAAGGTCCCTGCCAGCTC 1324

 367 roGlnAspAsnAspPheLeuSerArgLysAlaGlnAspCysTyrPheMet 383
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1325 CTCAAGACAATGATTTTTTATCTAGAAAAGCTCAAGACTGTTATTTTATG 1374

 384 LysLeuHisHisCysProGlyAsnHisSerTrpAspSerThrIleSerGl 400
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1375 AAGCTGCATCACTGTCCTGGAAATCACAGTTGGGATAGCACCATTTCTGG 1424

 400 ySerGlnArgAlaAlaPheCysAspHisLysThrThrProCysSerSerA 417
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1425 TTCTCAAAGGGCTGCATTCTGTGATCACAAGACCACTCCATGCTCTTCAG 1474

 417 laIleIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeuGlnPro 433
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1475 CAATAATAAATCCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAGCCT 1524

 434 GlyIleAlaGlnGlnTrpIleGlnSerLysArgGluAspIleValAsnGl 450
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1525 GGTATAGCCCAGCAGTGGATCCAGAGCAAAAGGGAAGACATTGTGAACCA 1574

 450 nMetThrGluAlaCysLeuAsnGlnSerLeuAspAlaLeuLeuSerArgA 467
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1575 AATGACAGAAGCCTGCCTTAACCAGTCGCTAGATGCCCTTCTGTCCAGGG 1624

 467 spLeuIleMetLysGluAspTyrGluLeuValSerThrLysProThrArg 483
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1625 ACTTGATCATGAAAGAGGACTATGAACTTGTTAGTACCAAGCCTACAAGG 1674

 484 ThrSerLysValArgGlnLeuLeuAspThrThrAspIleGlnGlyGluGl 500
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1675 ACCTCAAAAGTCAGACAATTACTAGACACTACTGACATCCAAGGAGAAGA 1724

 500 uPheAlaLysValIleValGlnLysLeuLysAspAsnLysGlnMetGlyL 517
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1725 ATTTGCCAAAGTTATAGTACAAAAATTGAAAGATAACAAACAAATGGGTC 1774

 517 euGlnProTyrProGluIleLeuValValSerArgSerProSerLeuAsn 533
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1775 TTCAGCCTTACCCGGAAATACTTGTGGTTTCTAGATCACCATCTTTAAAT 1824

 534 LeuLeuGlnAsnLysSerMet 540
     |||||||||||||||||||||
1825 TTACTTCAAAATAAAAGCATG 1845

seq_name: /cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:F21841

seq_documentation_block:
ID   F21841 standard; DNA; 463 BP.
XX
AC   F21841;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Human breast and ovarian cancer associated antigen gene SEQ ID 228.
XX
KW   Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW   nootropic; neurpprotective; antiviral; antiallergic; hepatotropic;
KW   antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;
KW   antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
KW   Addison's disease; allergy; autoimmune haemolytic anaemia;
KW   autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW   multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW   cardiovascular disorder; wound healing; neurological disease; ds.
XX
OS   Homo sapiens.
XX
PN   WO200055173-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05881.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-611515/58.
DR   P-PSDB; B58938.
XX
PT   New human breast and ovarian cancer associated gene sequences and the
PT   polypeptides encoded by these genes, useful in the prevention,
PT   treatment and diagnosis of cancer, immune disorders, cardiovascular
PT   disorders and neurological diseases -
XX
PS   Claim 1; Page 653; 1299pp; English.
XX
```

```
CC   Sequences F21614 - F22031 represent DNA sequences encoding human proteins
CC   B58711 - B59128. The DNA and protein sequences are associated with
CC   breast and ovarian cancer. Included in the invention are sequences
CC   F22032 - F22040 and B59129 which are used in the isolation and
CC   characterisation of the DNA and protein sequences of the invention. The
CC   breast and ovarian cancer associated DNA, protein, agonist or antagonist
CC   sequences exhibit cytostatic; immunosuppressive; nootropic;
CC   neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
CC   antiinflammatory; antiulcer; vulnerary; anticonvulsant; antibacterial;
CC   antifungal; antiparasitic and cardiant activity. The polynucleotide and
CC   protein sequences are used in the diagnosis of cancer, particularly
CC   breast and ovarian cancer. The nucleic acid sequences, proteins, agonists
CC   and agonists may also be used in the diagnosis, prevention and treatment
CC   of immune disorders e.g. Addison's disease, allergies, autoimmune
CC   haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC   disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC   cardiovascular disorders such as myocardial ischaemias; wound healing;
CC   neurological diseases such as cerebral anoxia and epilepsy; and
CC   infectious diseases.
XX
SQ   Sequence 463 BP; 150 A; 92 C; 89 G; 131 T; 1 other;


alignment_scores:
           Quality:  797.00            Length:    150
             Ratio:  5.349               Gaps:      0
 Percent Similarity:  99.333   Percent Identity: 99.333

alignment_block:
 US-09-445-223-1 x F21841    ..

 Align seg 1/1  to: F21841  from: 1  to: 463

      97 TyrMetProAsnGlySerLeuAsnGluLeuLeuHisArgLysThrGluTy 113
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       3 TACATGCCAAATGGATCATTAAATGAACTCCTACATAGGAAAACTGAATA 52

     113 rProAspValAlaTrpProLeuArgPheArgIleLeuHisGluIleAlaL 130
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      53 TCCTGATGTTGCTTGGCCATTGAGATTTCGCATCCTGCATGAAATTGCCC 102

     130 euGlyValAsnTyrLeuHisAsnMetThrProProLeuLeuHisHisAsp 146
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     103 TTGGTGTAAATTACCTGCACAATATGACTCCTCCTTTACTTCATCATGAC 152

     147 LeuLysThrGlnAsnIleLeuLeuAspAsnGluPheHisValLysIleAl 163
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     153 TTGAAGACTCAGAATATCTTATTGGACAATGAATTTCATGTTAAGATTGC 202

     163 aAspPheGlyLeuSerLysTrpArgMetMetSerLeuSerGlnSerArgS 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     203 AGATTTTGGTTTATCAAAGTGGCGCATGATGTCCCTCTCACAGTCACGAA 252

     180 erSerLysSerAlaProGluGlyGlyThrIleIleTyrMetProProGlu 196
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     253 GTAGCAAATCTGCACCAGAAGGAGGGACAATTATCTATATGCCACCTGAA 302

     197 AsnTyrGluProGlyGlnLysSerArgAlaSerIleLysHisAspIleTy 213
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     303 AACTATGAACCTGGACAAAAATCAAGGGCCAGTATCAAGCACGATATATA 352

     213 rSerTyrAlaValIleThrTrpGluValLeuSerArgLysGlnProPheG 230
         ||||||||||||||||||||||||||||   |||||||||||||||||||
     353 TAGCTATGCAGTTATCACATGGGAAGTGKTATCCAGAAAACAGCCTTTTG 402

     230 luAspValThrAsnProLeuGlnIleMetTyrSerValSerGlnGlyHis 246
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     403 AAGATGTCACCAATCCTTTGCAGATAATGTATAGTGTGTCACAAGGACAT 452

seq_name: /cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:Z61161

seq_documentation_block:
ID   Z61161 standard; DNA; 2370 BP.
```

```
XX
AC   Z61161;
XX
DT   30-MAY-2000  (first entry)
XX
DE   DNA encoding a death associated kinase with ankyrin repeats.
XX
KW   Death associated kinase protein containing ankyrin repeats; DAKAR;
KW   kinase; quality assurance agent; shelf life; marker; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             10..2370
FT                   /*tag= a
FT                   /transl_except= (pos:10..12, aa: Met)
XX
PN   WO200008177-A2.
XX
PD   17-FEB-2000.
XX
PF   04-AUG-1999;   99WO-US17576.
XX
PR   04-AUG-1998;   98US-0095269.
PR   11-SEP-1998;   98US-0099973.
PR   09-FEB-1999;   99US-0119353.
XX
PA   (IMMV ) IMMUNEX CORP.
XX
PI   Bird TA,  Virca GD;
XX
DR   WPI; 2000-195582/17.
DR   P-PSDB; Y69163.
XX
PT   Novel death associated kinase containing ankyrin repeats (DAKAR) used
PT   as molecular weight marker and as controls for peptide fragmentation  -
XX
PS   Claim 1; Page 9-10; 71pp; English.
XX
CC   The present sequence encodes a murine death associated kinase protein,
CC   containing ankyrin repeats, designated DAKAR. The DAKAR polynucleotides
CC   can be used to express the polypeptides, and as probes to identify
CC   nucleic acids encoding proteins having kinase activity. DAKAR
CC   polypeptides and fragmented polypeptides are used for purifying
CC   proteins, e.g. to measure protein activity; as quality assurance agents
CC   to monitor shelf life and stability of binding partner proteins; as
CC   research agents, e.g. in assays to determine protein kinase activity,
CC   to identify novel molecules involved in signal transduction pathways,
CC   and to identify therapeutic compounds which may interfere with
CC   apoptosis; as molecular weight and isoelectric focusing markers; as
CC   controls for peptide fragmentation; identification of unknown proteins,
CC   e.g. by comparison with proteins in databases; and for preparation of
CC   antibodies. The antibodies can be used in assays to detect the presence
CC   of the protein, and to purify the protein by immunoaffinity
CC   chromatography. The antibodies can also be used to block binding of
CC   the DAKAR polypeptides to their binding partners. Compounds that inhibit
CC   or enhance the kinase activity of DAKAR can be used to treat diseases
CC   characterized by overproduction or upregulated production or
CC   underproduction or downregulated production of DAKAR.
XX
SQ   Sequence 2370 BP; 536 A; 660 C; 700 G; 474 T; 0 other;


alignment_scores:
           Quality:  579.00            Length:    540
             Ratio:  1.969               Gaps:     26
 Percent Similarity:  54.444   Percent Identity: 34.074

alignment_block:
 US-09-445-223-1 x Z61161    ..

 Align seg 1/1  to: Z61161  from: 1  to: 2370
```

```
  10 LeuProThrIleProTyrHisLysLeuAlaAspLeuArgTyrLeuSerAr 26
     |||   |||:::         ::::::|||      :::   ::::::
  49 CTGCGCACCTTCGACGCCGGCGAATTCGCAGGCTGGGAGAAGGTGGGCTC 98

  26 gGlyAlaSerGlyThrValSerSerAlaArgHisAlaAspTrpArgValG 43
      |||:::   |||   |||   ::::::||||||:::   |||::::::
  99 GGGCGGCTTCGGGCAGGTGTACAAGGTGCGCCATGTGCACTGGAAGACGT 148

  43 lnValAlaValLys.........HisLeuHisIleHisThrProLeuLeu 56
       :::|||:::|||            ||||||:::
 149 GGCTCGCGATCAAGTGCTCGCCCAGTCTGCACGTCGAC............ 186

  57 AspSerGluArgLysAspValLeuArgGluAlaGluIleLeuHisLysAl 73
     |||   ||||||   ::::::|||:::||||||:::   ::::::   ||
 187 GACAGGGAACGAATGGAGCTCCTGGAGGAAGCTAAGAAGATGGAGATGGC 236

  73 aArgPheSerTyrIlePheProIleLeuGlyIleCysAsnGluProGluP 90
     |:::|||   ||||||:::|||:::   |||||||||:::||||||
 237 CAAGTTCCGATACATTCTACCTGTGTACGGCATATGCCAGGAACCT.... 282

  90 heLeuGlyIleValThrGluTyrMetProAsnGlySerLeuAsnGluLeu 106
       :::|||:::|||   |||||||||   :::|||||||||::::::|||
 283 ..GTCGGCTTGGTCATGGAGTACATGGAGACAGGCTCCCTGGAGAAGCTG 330

 107 LeuHisArgLysThrGluTyrProAspValAlaTrpProLeuArgPheAr 123
     |||      :::               :::   |||   |||||||||||
 331 CTGGCCTCAGAG...............CCATTGCCTTGGGACCTGCGCTTTCG 368

 123 gIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisAsnMetThrP 140
     ||||:::||||||   |||:::|||:::|||:::||||||   |||:::|
 369 CATCGTGCACGAGACAGCCGTGGGCATGAACTTCCTGCATTGCATGTCTC 418

 140 roProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeuAspAsn 156
     ||||||||||||||   |||||||||      |||||||||||||||
 419 CGCCACTGCTGCACCTAGACCTGAAGCCAGCGAACATCCTGCTGGATGCC 468

 157 GluPheHisValLysIleAlaAspPheGlyLeuSerLysTrpArgMetMe 173
     ::::::||||||||||||:::||||||||||||:::|||   :::   ||
 469 CACTACCATGTCAAGATTTCTGACTTTGGGCTGGCCAAGTGCAATGGCAT 518

 173 tSerLeuSerGlnSerArgSerSerLysSerAlaProGluGly......G 188
     ||||   |||::::::   |||               :::|||      |
 519 GTCCCACTCTCATGACCTCAGCATG............GATGGCCTGTTTG 556

 188 lyThrIleIleTyrMetProProGluAsnTyrGluProGlyGlnLysSer 204
     ||||||||   |||:::|||||||||:::   :::      :::||||||
 557 GTACAATCGCTTACCTCCCTCCAGAGCGAATTCGT......GAGAAGAGC 600

 205 ArgAla...SerIleLysHisAspIleTyrSerTyrAlaValIleThrTr 220
     |||      :::   |||||||||:::||||||:::|||::::::   ||
 601 CGCTTGTTTGACACCAAACATGATGTATACAGCTTCGCCATTGTGATCTG 650

 220 pGluValLeuSerArgLysGlnProPheGluAspValThrAsnProLeuG 237
     |   ||||||::::::|||:::||||||   |||      |||   |||:
 651 GGGTGTGCTTACACAGAAGAAGCCATTTGCAGATGAAAAGAACATCCTAC 700

 237 lnIleMetTyrSerValSerGlnGlyHisArgProValIleAsnGluGlu 253
     ::||||||   :::|||   :::||||||||||||            |||
 701 ACATCATGATGAAAGTGGTAAAGGGCCACCGCCCA............GAG 738

 254 SerLeuProTyrAspIleProHisArg.........AlaArgMetIleSe 267
           |||         |||:::            |||   :::|||::
 739 CTGCCACCCATCTGCAGACCCCGGCCGCGTGCCTGTGCCAGCCTGATAGG 788

 267 rLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerPheL 284
     :|||::::::      |||      :::|||:::   ||||||
 789 GCTCATGCAACGGTGCTGGCATGCAGACCCACAGGTGCGGCCC....... 831

 284 euLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGluIle 300
                                        ||||||:::||||||
 832 ...................................ACCTTCCAAGAAATT 846

 301 ThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSerVa 317
     |||
 847 ACC............................................... 849

 317 lSerSerAlaIleHisLeuCysAspLysLysLysMetGluLeuSerLeuA 334
      |||::::::      ||||||:::|||         |||::::::
 850 .TCTGAAACAGAAGACCTTTGTGAGAAGCCTGATGAGGAGGTGAAA.... 894

 334 snIleProValAsnHisGlyProGlnGluGluSerCysGlySerSerGln 350
             :::   |||   |||   |||:::|||      :::|||:::
 895 .....GACCTGGCTCATGAGCCAGGCGAGAAAAGCTCTCTAGAGTCCAAG 939

 351 LeuHisGluAsnSerGlySerProGluThrSerArgSerLeuProAlaPr 367
        :::   :::      |||   :::   :::|||:::         ||
 940 AGTGAGGCCAGGCCCGAGTCCTCACGCCTCAAGCGCGCCTCTGCTCCCCC 989

 367 oGlnAspAsnAsp.PheLeuSerArgLysAlaGlnAspCysTyrPheMet 383
     |   |||||||||    :::|||            :::|||
 990 CTTCGATAACGACTGCAGTCTCT.........CCGAGTTGC......... 1021

 384 LysLeuHisHisCysProGlyAsnHisSerTrpAspSerThrIleSerGl 400
                 |||         |||||||||               ||
1022 ............TGT.........CACAGTTGG.........ACTCTGGG 1041

 400 ySerGlnArgAlaAlaPheCysAspHisLysThrThrProCysSerSerA 417
     ||||   |||                        :::|||   |||||||
1042 ATCTCCCAGACTCTTG...............AAGGCCCCGAAGAGCTCAG 1076

 417 laIleIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeuGlnPro 433
     ||   :::   ||||||            |||::::::      :::|||
1077 CCGAAGTT...CCTCTG............AATGCAAGCTCCCATCGTCCA 1111

 434 GlyIleAlaGln..................GlnTrpIleGlnSerLysAr 444
     ::::::|||:::                  ||||||   |||
1112 GCAGTGGCAAGAGGCTCTCGGGGGTGTCCTCAGTGGACTCAGCCTTTTCC 1161

 444 gGluAspIleValAsnGlnMetThrGluAlaCysLeuAsnGlnSerLeuA 461
         :::                  :::   |||
1162 TCCAGAG..................GATCGCTGT...............C 1178

 461 spAlaLeuLeuSerArgAspLeuIleMetLysGluAspTyrGluLeuVal 477
       :::|||||||||
1179 ACTGTCTTTTGAGC.................................... 1192

 478 SerThrLysProThrArgThrSerLysValArgGlnLeu..LeuAspThr 493
                          :::|||::::::|||     |||
1193 .....................GGGAAGCTTCAACAGGCGACCTGGCCCC 1221

 494 ThrAspIleGlnGlyGluGluPheAlaLysValIleValGlnLysLeuLy 510
     ||||||||||||   ::::::::::::   :::|||::::::
1222 ACAGACATCCAGAAGAAGAAGCTAGTGGATGCCATCATATCAGGGGACAC 1271

 510 sAspAsnLysGlnMetGlyLeuGlnProTyrProGluIleLeuValValS 527
      ::::::   :::      |||||||||            ||||||::::
1272 CAGCAGGCTGATGAAGATCCTACAGCCCCAAGATGTGGACTTGGTTCTAG 1321

 527 erArgSerProSerLeu 532
     ::   |||   ||||||
1322 ACAGCAGTGCCAGCCTG 1338

seq_name: /cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:Z61784

seq_documentation_block:
ID   Z61784 standard; cDNA; 3516 BP.
XX
AC   Z61784;
XX
DT   27-MAR-2000  (first entry)
XX
DE   cDNA encoding murine protein kinase/ankyrin homologue, SEQ ID NO:257.
XX
KW   Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
KW   embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
KW   secreted; transmembrane; inflammation; cancer; neurological disease;
```

```
KW   angiogenesis; tumour vascularisation; growth disorder;
KW   developmental disorder; skin wound; hair follicle disorder;
KW   anti-inflammatory; cytostatic; neuroprotective; vulnery; ss.
XX
OS   Mus sp.
XX
PN   WO9955865-A1.
XX
PD   04-NOV-1999.
XX
PF   29-APR-1999;   99WO-NZ00051.
XX
PR   29-APR-1998;   98US-0069726.
PR   09-NOV-1998;   98US-0188930.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Strachan L,  Sleeman M,  Watson JD,  Onrust R,  Kumble A,  Murison JG;
XX
DR   WPI; 2000-072177/06.
DR   P-PSDB; Y76079.
XX
PT   Novel polynucleotides useful for the treatment of various conditions
PT   including wounds and cancer -
XX
PS   Claim 1; Page 158-159; 235pp; English.
XX
CC   The invention relates to novel nucleic acid sequences derived from rat
CC   dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
CC   and mouse embryonic skin, keratinocyte stem cells and transit amplifying
CC   cells. Polypeptides of the invention may be used to treat inflammation,
CC   cancer and neurological diseases. The proteins may be used to stimulate
CC   the growth and motility of keratinocytes, to inhibit the growth of
CC   cancer cells, to modulate angiogenesis and tumour vascularisation, to
CC   modulate skin inflammation, to modulate epithelial cell growth and to
CC   inhibit binding of HIV-1 to leukocytes. The invention may also be used to
CC   treat growth and developmental defects, skin wounds and hair follicle
CC   disorders. Sequences Z61606-Z61832 represent cDNA sequences derived from
CC   several mouse, rat or human skin cell types. Sequences Z61606-Z61649,
CC   Z61725-Z61765, Z61802-Z61811 and Z61826 encode proteins with an
CC   N-terminal signal sequence, indicating that the proteins are secreted.
CC   Sequences Z61650-Z61668, Z61766-Z61780, Z61812-Z61817 and Z61827-Z61829
CC   encode proteins with one or more putative transmembrane domains.
XX
SQ   Sequence 3516 BP; 810 A; 886 C; 988 G; 832 T; 0 other;


alignment_scores:
           Quality:  579.00             Length:    540
             Ratio:  1.969                Gaps:     26
 Percent Similarity:  54.444   Percent Identity: 34.074

alignment_block:
 US-09-445-223-1 x Z61784   ..

 Align seg 1/1  to: Z61784  from: 1  to: 3516

      10 LeuProThrIleProTyrHisLysLeuAlaAspLeuArgTyrLeuSerAr 26
         |||   |||:::         ::::::|||      :::   ::::::
      43 CTGCGCACCTTCGACGCCGGCGAATTCGCAGGCTGGGAGAAGGTGGGCTC 92

      26 gGlyAlaSerGlyThrValSerSerAlaArgHisAlaAspTrpArgValG 43
          |||:::   |||   |||   ::::::||||||:::   |||::::::
      93 GGGCGGCTTCGGGCAGGTGTACAAGGTGCGCCATGTGCACTGGAAGACGT 142

      43 lnValAlaValLys.........HisLeuHisIleHisThrProLeuLeu 56
           :::|||:::|||            ||||||:::
     143 GGCTCGCGATCAAGTGCTCGCCCAGTCTGCACGTCGAC............ 180

      57 AspSerGluArgLysAspValLeuArgGluAlaGluIleLeuHisLysAl 73
         |||   ||||||   ::::::|||:::||||||:::   ::::::   ||
     181 GACAGGGAACGAATGGAGCTCCTGGAGGAAGCTAAGAAGATGGAGATGGC 230

      73 aArgPheSerTyrIlePheProIleLeuGlyIleCysAsnGluProGluP 90
         |:::|||   ||||||:::|||:::   |||||||||:::||||||
     231 CAAGTTCCGATACATTCTACCTGTGTACGGCATATGCCAGGAACCT.... 276

      90 heLeuGlyIleValThrGluTyrMetProAsnGlySerLeuAsnGluLeu 106
           :::|||:::|||   |||||||||   :::|||||||||::::::|||
     277 ..GTCGGCTTGGTCATGGAGTACATGGAGACAGGCTCCCTGGAGAAGCTG 324

     107 LeuHisArgLysThrGluTyrProAspValAlaTrpProLeuArgPheAr 123
         |||      :::               :::   |||   |||||||||||
     325 CTGGCCTCAGAG...........CCATTGCCTTGGGACCTGCGCTTTCG 362

     123 gIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisAsnMetThrP 140
         ||||:::||||||   |||:::|||:::|||:::||||||   |||:::|
     363 CATCGTGCACGAGACAGCCGTGGGCATGAACTTCCTGCATTGCATGTCTC 412

     140 roProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeuAspAsn 156
         ||||||||||||||   |||||||||      |||||||||||||||
     413 CGCCACTGCTGCACCTAGACCTGAAGCCAGCGAACATCCTGCTGGATGCC 462

     157 GluPheHisValLysIleAlaAspPheGlyLeuSerLysTrpArgMetMe 173
         ::::::||||||||||||:::||||||||||||:::|||   :::   ||
     463 CACTACCATGTCAAGATTTCTGACTTTGGGCTGGCCAAGTGCAATGGCAT 512

     173 tSerLeuSerGlnSerArgSerSerLysSerAlaProGluGly......G 188
         ||||   |||::::::   |||               :::|||      |
     513 GTCCCACTCTCATGACCTCAGCATG...........GATGGCCTGTTTG 550

     188 lyThrIleIleTyrMetProProGluAsnTyrGluProGlyGlnLysSer 204
         ||||||||   |||:::|||||||||:::   :::      :::||||||
     551 GTACAATCGCTTACCTCCCTCCAGAGCGAATTCGT......GAGAAGAGC 594

     205 ArgAla...SerIleLysHisAspIleTyrSerTyrAlaValIleThrTr 220
         |||      :::   |||||||||:::||||||:::|||::::::   ||
     595 CGCTTGTTTGACACCAAACATGATGTATACAGCTTCGCCATTGTGATCTG 644

     220 pGluValLeuSerArgLysGlnProPheGluAspValThrAsnProLeuG 237
         |   ||||||::::::|||:::||||||   |||      |||   |||:
     645 GGGTGTGCTTACACAGAAGAAGCCATTTGCAGATGAAAAGAACATCCTAC 694

     237 lnIleMetTyrSerValSerGlnGlyHisArgProValIleAsnGluGlu 253
         ::||||||   :::|||   :::||||||||||||            |||
     695 ACATCATGATGAAAGTGGTAAAGGGCCACCGCCCA............GAG 732

     254 SerLeuProTyrAspIleProHisArg.........AlaArgMetIleSe 267
               |||         |||:::            |||   :::|||::
     733 CTGCCACCCATCTGCAGACCCCGGCCGCGTGCCTGTGCCAGCCTGATAGG 782

     267 rLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerPheL 284
         :|||::::::      |||      :::|||:::   ||||||
     783 GCTCATGCAACGGTGCTGGCATGCAGACCCACAGGTGCGGCCC....... 825

     284 euLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGluIle 300
                                            ||||||:::||||||
     826 ...................................ACCTTCCAAGAAATT 840

     301 ThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSerVa 317
         |||
     841 ACC............................................... 843

     317 lSerSerAlaIleHisLeuCysAspLysLysLysMetGluLeuSerLeuA 334
          |||::::::      ||||||:::|||         |||::::::
     844 .TCTGAAACAGAAGACCTTTGTGAGAAGCCTGATGAGGAGGTGAAA.... 888

     334 snIleProValAsnHisGlyProGlnGluGluSerCysGlyGlySerGln 350
                 :::   |||   |||   |||:::|||      :::|||:::
     889 .....GACCTGGCTCATGAGCCAGGCGAGAAAAGCTCTCTAGAGTCCAAG 933

     351 LeuHisGluAsnSerGlySerProGluThrSerArgSerLeuProAlaPr 367
            :::   :::      |||   :::   :::|||:::         ||
     934 AGTGAGGCCAGGCCCGAGTCCTCACGCCTCAAGCGCGCCTCTGCTCCCCC 983

     367 oGlnAspAsnAsp.PheLeuSerArgLysAlaGlnAspCysTyrPheMet 383
```

```
     |   |||||||||    :::|||            :::|||
 984 CTTCGATAACGACTGCAGTCTCT.........CCGAGTTGC.......... 1015

 384 LysLeuHisHisCysProGlyAsnHisSerTrpAspSerThrIleSerGl 400
                 |||         |||||||||               ||
1016 ............TGT.........CACAGTTGG.........ACTCTGGG 1035

 400 ySerGlnArgAlaAlaPheCysAspHisLysThrThrProCysSerSerA 417
     ||||   |||                         :::|||  |||||||
1036 ATCTCCCAGACTCTTG...............AAGGCCCCGAAGAGCTCAG 1070

 417 laIleIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeuGlnPro 433
     ||   :::   ||||||            |||::::::      :::|||
1071 CCGAAGTT...CCTCTG............AATGCAAGCTCCCATCGTCCA 1105

 434 GlyIleAlaGln..................GlnTrpIleGlnSerLysAr 444
     ::::::|||:::                  ||||||   |||
1106 GCAGTGGCAAGAGGCTCTCGGGGGTGTCCTCAGTGGACTCAGCCTTTTCC 1155

 444 gGluAspIleValAsnGlnMetThrGluAlaCysLeuAsnGlnSerLeuA 461
         :::                  :::   |||
1156 TCCAGAG..................GATCGCTGT...............C 1172

 461 spAlaLeuLeuSerArgAspLeuIleMetLysGluAspTyrGluLeuVal 477
       :::|||||||||
1173 ACTGTCTTTTGAGC.................................... 1186

 478 SerThrLysProThrArgThrSerLysValArgGlnLeu..LeuAspThr 493
                          :::|||::::::|||     |||
1187 .....................GGGAAGCTTCAACAGGCGACCTGGGCCCC 1215

 494 ThrAspIleGlnGlyGluGluPheAlaLysValIleValGlnLysLeuLy 510
     ||||||||||||   ::::::::::::   :::|||::::::
1216 ACAGACATCCAGAAGAAGAAGCTAGTGGATGCCATCATATCAGGGGACAC 1265

 510 sAspAsnLysGlnMetGlyLeuGlnProTyrProGluIleLeuValValS 527
      ::::::   :::      |||||||||            ||||||::::
1266 CAGCAGGCTGATGAAGATCCTACAGCCCCAAGATGTGGACTTGGTTCTAG 1315

 527 erArgSerProSerLeu 532
     ::   |||   ||||||
1316 ACAGCAGTGCCAGCCTG 1332

seq_name: /cgn1_8/gcgdata/geneseq/geneseqn/NA2001.DAT:C99717

seq_documentation_block:
ID   C99717 standard; cDNA; 3516 BP.
XX
AC   C99717;
XX
DT   08-MAR-2001  (first entry)
XX
DE   Skin cell cDNA, SEQ ID NO: 257.
XX
KW   Mouse; skin cell; cytostatic; antiinflammatory; anti-HIV;
KW   nootropic; neuroprotective; vulnerary; immunomodulatory; vaccine;
KW   keratinocyte growth stimulation; cancer; angiogenesis inhibition;
KW   inflammation; neurological disease; ss.
XX
OS   Mus sp.
XX
PN   WO200069884-A2.
XX
PD   23-NOV-2000.
XX
PF   15-MAY-2000; 2000WO-NZ00075.
XX
PR   14-MAY-1999;   99US-0312283.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Watson JD,  Strachan L,  Onrust R,  Sleeman M,  Kumble KD,  Murison JG;
XX
```

```
DR   WPI; 2001-007495/01.
DR   P-PSDB; B55958.
XX
PT   New isolated polynucleotide used in the identification of genetic
PT   disorders and encoding polypeptides used for treating inflammatory
PT   disease, cancer and neurological diseases -
XX
PS   Claim 1; Page 216-218; 352pp; English.
XX
CC   The present polynucleotide encodes a polypeptide which is expressed in
CC   mammalian skin cells. The polypeptide is useful for stimulating
CC   keratinocyte growth and motility, inhibiting the growth of cancer cells,
CC   modulating angiogenesis, inhibiting angiogenesis and vascularisation of
CC   tumours, modulating skin inflammation, stimulating the growth of
CC   epithelial cells, inhibiting the binding of human immunodeficiency virus
CC   (HIV)-1 to leukocytes, and treating inflammmatory disease, cancer and
CC   neurological diseases. The polynucleotide can be used as a marker, in
CC   the identification of genetic disorders, and for the design of
CC   oligonucleotides for examining expression patterns.
XX
SQ   Sequence 3516 BP; 810 A; 886 C; 988 G; 832 T; 0 other;

alignment_scores:
          Quality:  579.00            Length:    540
            Ratio:  1.969               Gaps:     26
 Percent Similarity:  54.444   Percent Identity: 34.074

alignment_block:
 US-09-445-223-1 x C99717   ..

 Align seg 1/1  to: C99717  from: 1  to: 3516

  10 LeuProThrIleProTyrHisLysLeuAlaAspLeuArgTyrLeuSerAr  26
     |||   |||:::         ::::::|||      :::   ::::::
  43 CTGCGCACCTTCGACGCCGGCGAATTCGCAGGCTGGGAGAAGGTGGGCTC  92

  26 gGlyAlaSerGlyThrValSerSerAlaArgHisAlaAspTrpArgValG  43
      |||:::   |||   |||   ::::::||||||:::   |||::::::
  93 GGGCGGCTTCGGGCAGGTGTACAAGGTGCGCCATGTGCACTGGAAGACGT 142

  43 lnValAlaValLys.........HisLeuHisIleHisThrProLeuLeu  56
       :::|||:::|||            ||||||:::
 143 GGCTCGCGATCAAGTGCTCGCCCAGTCTGCACGTCGAC............ 180

  57 AspSerGluArgLysAspValLeuArgGluAlaGluIleLeuHisLysAl  73
     |||   ||||||   ::::::|||:::||||||:::   ::::::   ||
 181 GACAGGGAACGAATGGAGCTCCTGGAGGAAGCTAAGAAGATGGAGATGGC 230

  73 aArgPheSerTyrIlePheProIleLeuGlyIleCysAsnGluProGluP  90
     |:::|||   ||||||:::|||:::   |||||||||:::||||||
 231 CAAGTTCCGATACATTCTACCTGTGTACGGCATATGCCAGGAACCT.... 276

  90 heLeuGlyIleValThrGluTyrMetProAsnGlySerLeuAsnGluLeu 106
       :::|||:::|||   |||||||||   :::|||||||||::::::|||
 277 ..GTCGGCTTGGTCATGGAGTACATGGAGACAGGCTCCCTGGAGAAGCTG 324

 107 LeuHisArgLysThrGluTyrProAspValAlaTrpProLeuArgPheAr 123
     |||      :::               :::   |||   |||||||||||
 325 CTGGCCTCAGAG............CCATTGCCTTGGGACCTGCGCTTTCG 362

 123 gIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisAsnMetThrP 140
     ||||:::||||||   |||:::|||:::|||:::||||||   |||:::|
 363 CATCGTGCACGAGACAGCCGTGGGCATGAACTTCCTGCATTGCATGTCTC 412

 140 roProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeuAspAsn 156
     ||||||||||||||   |||||||||      |||||||||||||||
 413 CGCCACTGCTGCACCTAGACCTGAAGCCAGCGAACATCCTGCTGGATGCC 462

 157 GluPheHisValLysIleAlaAspPheGlyLeuSerLysTrpArgMetMe 173
     ::::::||||||||||||:::||||||||||||:::|||   :::   ||
 463 CACTACCATGTCAAGATTTCTGACTTTGGGCTGGCCAAGTGCAATGGCAT 512
```

```
 173 tSerLeuSerGlnSerArgSerSerLysSerAlaProGluGly......G 188
     ||||   |||::::::    |||               :::|||      |
 513 GTCCCACTCTCATGACCTCAGCATG.............GATGGCCTGTTTG 550

 188 lyThrIleIleTyrMetProProGluAsnTyrGluProGlyGlnLysSer 204
     ||||||||   |||:::|||||||||:::    :::      :::||||||
 551 GTACAATCGCTTACCTCCCTCCAGCGAATTCGT......GAGAAGAGC 594

 205 ArgAla...SerIleLysHisAspIleTyrSerTyrAlaValIleThrTr 220
     |||      :::   |||||||||:::||||||:::|||::::::    ||
 595 CGCTTGTTTGACACCAAACATGATGTATACAGCTTCGCCATTGTGATCTG 644

 220 pGluValLeuSerArgLysGlnProPheGluAspValThrAsnProLeuG 237
     |   ||||||::::::|||:::||||||    |||      |||   |||:
 645 GGGTGTGCTTACACAGAAGAAGCCATTTGCAGATGAAAAGAACATCCTAC 694

 237 lnIleMetTyrSerValSerGlnGlyHisArgProValIleAsnGluGlu 253
     ::||||||   :::|||   :::|||||||||||||            |||
 695 ACATCATGATGAAAGTGGTAAAGGGCCACCGCCCA.............GAG 732

 254 SerLeuProTyrAspIleProHisArg..........AlaArgMetIleSe 267
           |||         |||:::             |||   :::|||::
 733 CTGCCACCCATCTGCAGACCCCGGCCGCGTGCCTGTGCCAGCCTGATAGG 782

 267 rLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerPheL 284
     :|||::::::      |||      :::|||:::    ||||||
 783 GCTCATGCAACGGTGCTGGCATGCAGACCCACAGGTGCGGCCC........ 825

 284 euLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGluIle 300
                                         ||||||:::||||||
 826 ....................................ACCTTCCAAGAAATT 840

 301 ThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSerVa 317
     |||
 841 ACC................................................ 843

 317 lSerSerAlaIleHisLeuCysAspLysLysLysMetGluLeuSerLeuA 334
      |||::::::      ||||||:::|||          |||::::::
 844 .TCTGAAACAGAAGACCTTTGTGAGAAGCCTGATGAGGAGGTGAAA.... 888

 334 snIleProValAsnHisGlyProGlnGluGluSerCysGlySerSerGln 350
             :::   |||   |||   |||:::|||       :::|||:::
 889 .....GACCTGGCTCATGAGCCAGGCGAGAAAAGCTCTCTAGAGTCCAAG 933

 351 LeuHisGluAsnSerGlySerProGluThrSerArgSerLeuProAlaPr 367
        :::   :::      |||   :::    :::|||:::         ||
 934 AGTGAGGCCAGGCCCGAGTCCTCACGCCTCAAGCGCGCCTCTGCTCCCCC 983

 367 oGlnAspAsnAsp.PheLeuSerArgLysAlaGlnAspCysTyrPheMet 383
     |   |||||||||    :::|||             :::|||
 984 CTTCGATAACGACTGCAGTCTCT.........CCGAGTTGC.......... 1015

 384 LysLeuHisHisCysProGlyAsnHisSerTrpAspSerThrIleSerGl 400
                 |||         |||||||||                ||
1016 ............TGT.........CACAGTTGG.........ACTCTGGG 1035

 400 ySerGlnArgAlaAlaPheCysAspHisLysThrThrProCysSerSerA 417
     ||||   |||                         :::|||   |||||||
1036 ATCTCCCAGACTCTTG................AAGGCCCCGAAGAGCTCAG 1070

 417 laIleIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeuGlnPro 433
     ||   :::   ||||||             |||::::::      :::|||
1071 CCGAAGTT...CCTCTG.............AATGCAAGCTCCCATCGTCCA 1105

 434 GlyIleAlaGln...................GlnTrpIleGlnSerLysAr 444
     ::::::|||:::                   ||||||   |||
1106 GCAGTGGCAAGAGGCTCTCGGGGGTGTCCTCAGTGGACTCAGCCTTTTCC 1155

 444 gGluAspIleValAsnGlnMetThrGluAlaCysLeuAsnGlnSerLeuA 461
         :::                  :::    |||
1156 TCCAGAG..................GATCGCTGT................C 1172

 461 spAlaLeuLeuSerArgAspLeuIleMetLysGluAspTyrGluLeuVal 477
       :::|||||||||
1173 ACTGTCTTTTGAGC..................................... 1186

 478 SerThrLysProThrArgThrSerLysValArgGlnLeu..LeuAspThr 493
                          :::|||::::::|||      |||
1187 .....................GGGAAGCTTCAACAGGCGACCTGGGCCCC 1215

 494 ThrAspIleGlnGlyGluGluPheAlaLysValIleValGlnLysLeuLy 510
     ||||||||||||   ::::::::::::    :::|||::::::
1216 ACAGACATCCAGAAGAAGAAGCTAGTGGATGCCATCATATCAGGGGACAC 1265

 510 sAspAsnLysGlnMetGlyLeuGlnProTyrProGluIleLeuValValS 527
      ::::::   :::      |||||||||             ||||||::::
1266 CAGCAGGCTGATGAAGATCCTACAGCCCCAAGATGTGGACTTGGTTCTAG 1315

 527 erArgSerProSerLeu 532
     ::   |||   ||||||
1316 ACAGCAGTGCCAGCCTG 1332

seq_name: /cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:Z61830

seq_documentation_block:
ID   Z61830 standard; cDNA; 1774 BP.
XX
AC   Z61830;
XX
DT   27-MAR-2000  (first entry)
XX
DE   cDNA encoding murine RIP protein kinase homologue, SEQ ID NO:403.
XX
KW   Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
KW   embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
KW   secreted; transmembrane; inflammation; cancer; neurological disease;
KW   angiogenesis; tumour vascularisation; growth disorder;
KW   developmental disorder; skin wound; hair follicle disorder;
KW   anti-inflammatory; cytostatic; neuroprotective; vulnery; ss.
XX
OS   Mus sp.
XX
PN   WO9955865-A1.
XX
PD   04-NOV-1999.
XX
PF   29-APR-1999;   99WO-NZ00051.
XX
PR   29-APR-1998;   98US-0069726.
PR   09-NOV-1998;   98US-0188930.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Strachan L,  Sleeman M,  Watson JD,  Onrust R,  Kumble A,  Murison JG;
XX
DR   WPI; 2000-072177/06.
DR   P-PSDB; Y76123.
XX
PT   Novel polynucleotides useful for the treatment of various conditions
PT   including wounds and cancer -
XX
PS   Claim 1; Page 228-229; 235pp; English.
XX
CC   The invention relates to novel nucleic acid sequences derived from rat
CC   dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
CC   and mouse embryonic skin, keratinocyte stem cells and transit amplifying
CC   cells. Polypeptides of the invention may be used to treat inflammation,
CC   cancer and neurological diseases. The proteins may be used to stimulate
CC   the growth and motility of keratinocytes, to inhibit the growth of
CC   cancer cells, to modulate angiogenesis and tumour vascularisation, to
CC   modulate skin inflammation, to modulate epithelial cell growth and to
CC   inhibit binding of HIV-1 to leukocytes. The invention may also be used to
CC   treat growth and developmental defects, skin wounds and hair follicle
CC   disorders. Sequences Z61606-Z61832 represent cDNA sequences derived from
CC   several mouse, rat or human skin cell types. Sequences Z61606-Z61649,
CC   Z61725-Z61765, Z61802-Z61811 and Z61826 encode proteins with an
```

```
CC   N-terminal signal sequence, indicating that the proteins are secreted.
CC   Sequences Z61650-Z61668, Z61766-Z61780, Z61812-Z61817 and Z61827-Z61829
CC   encode proteins with one or more putative transmembrane domains.
XX
SQ   Sequence 1774 BP; 406 A; 490 C; 526 G; 352 T; 0 other;

alignment_scores:
           Quality:  577.00             Length:     540
             Ratio:  1.963                Gaps:      26
 Percent Similarity:  54.444   Percent Identity: 33.889

alignment_block:
 US-09-445-223-1 x Z61830    ..

 Align seg 1/1  to: Z61830  from: 1  to: 1774

  10 LeuProThrIleProTyrHisLysLeuAlaAspLeuArgTyrLeuSerAr 26
     |||   |||:::         ::::::|||      :::   ::::::
  45 CTGCGCACCTTCGACGCCGGCGAATTCGCAGGCTGGGAGAAGGTGGGCTC 94

  26 gGlyAlaSerGlyThrValSerSerAlaArgHisAlaAspTrpArgValG 43
      |||:::   |||   |||   ::::::||||||:::   |||::::::
  95 GGGCGGCTTCGGGCAGGTGTACAAGGTGCGCCATGTGCACTGGAAGACGT 144

  43 lnValAlaValLys.........HisLeuHisIleHisThrProLeuLeu 56
       :::|||:::|||            ||||||:::
 145 GGCTCGCGATCAAGTGCTCGCCCAGTCTGCACGTCGAC............ 182

  57 AspSerGluArgLysAspValLeuArgGluAlaGluIleLeuHisLysAl 73
     |||   ||||||   ::::::|||:::||||||:::   ::::::    ||
 183 GACAGGGAACGAATGGAGCTCCTGGAGGAAGCTAAGAAGATGGAGATGGC 232

  73 aArgPheSerTyrIlePheProIleLeuGlyIleCysAsnGluProGluP 90
     |:::|||   ||||||:::|||:::   |||||||||:::||||||
 233 CAAGTTCCGATACATTCTACCTGTGTACGGCATATGCCAGGAACCT.... 278

  90 heLeuGlyIleValThrGluTyrMetProAsnGlySerLeuAsnGluLeu 106
       :::|||:::|||   |||||||||   :::|||||||||::::::|||
 279 ..GTCGGCTTGGTCATGGAGTACATGGAGACAGGCTCCCTGGAGAAGCTG 326

 107 LeuHisArgLysThrGluTyrProAspValAlaTrpProLeuArgPheAr 123
     |||      :::               :::   |||   |||||||||||
 327 CTGGCCTCAGAG............CCATTGCCTTGGGACCTGCGCTTTCG 364

 123 gIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisAsnMetThrP 140
     ||||:::||||||   |||:::|||:::|||:::||||||   |||:::|
 365 CATCGTGCACGAGACAGCCGTGGGCATGAACTTCCTGCATTGCATGTCTC 414

 140 roProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeuAspAsn 156
     ||||||||||||||   |||||||||      |||||||||||||||
 415 CGCCACTGCTGCACCTAGACCTGAAGCCAGCGAACATCCTGCTGGATGCC 464

 157 GluPheHisValLysIleAlaAspPheGlyLeuSerLysTrpArgMetMe 173
     ::::::||||||||||||:::||||||||||||:::|||   :::   ||
 465 CACTACCATGTCAAGATTTCTGACTTTGGGCTGGCCAAGTGCAATGGCAT 514

 173 tSerLeuSerGlnSerArgSerSerLysSerAlaProGluGly......G 188
     ||||   |||::::::   |||               :::|||      |
 515 GTCCCACTCTCATGACCTCAGCATG............GATGGCCTGTTTG 552

 188 lyThrIleIleTyrMetProProGluAsnTyrGluProGlyGlnLysSer 204
     ||||||||   |||:::|||||||||:::   :::      :::||||||
 553 GTACAATCGCTTACCTCCCTCCAGAGCGAATTCGT......GAGAAGAGC 596

 205 ArgAla...SerIleLysHisAspIleTyrSerTyrAlaValIleThrTr 220
     |||      :::   |||||||||:::||||||:::|||::::::   ||
 597 CGCTTGTTTGACACCAAACATGATGTATACAGCTTCGCCATTGTGATCTG 646

 220 pGluValLeuSerArgLysGlnProPheGluAspValThrAsnProLeuG 237
     |   ||||||::::::|||:::||||||   |||      |||   |||:
 647 GGGTGTGCTTACACAGAAGAAGCCATTTGCAGATGAAAAGAACATCCTAC 696

 237 lnIleMetTyrSerValSerGlnGlyHisArgProValIleAsnGluGlu 253
     ::||||||   :::|||   :::||||||||||||            |||
 697 ACATCATGATGAAAGTGGTAAAGGGCCACCGCCCA............GAG 734

 254 SerLeuProTyrAspIleProHisArg.........AlaArgMetIleSe 267
           |||         |||:::            |||   :::|||::
 735 CTGCCACCCATCTGCAGACCCCGGCCGCGTGCCTGTGCCAGCCTGATAGG 784

 267 rLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerPheL 284
     ::::::::::      |||      :::|||:::   ||||||
 785 GATAATGCAACGGTGCTGGCATGCAGACCCACAGGTGCGGCCC....... 827

 284 euLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGluIle 300
                                        ||||||:::||||||
 828 ...................................ACCTTCCAAGAAATT 842

 301 ThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSerVa 317
     |||
 843 ACC............................................... 845

 317 lSerSerAlaIleHisLeuCysAspLysLysLysMetGluLeuSerLeuA 334
      |||::::::      ||||||:::|||         |||::::::
 846 .TCTGAAACAGAAGACCTTTGTGAGAAGCCTGATGAGGAGGTGAAA.... 890

 334 snIleProValAsnHisGlyProGlnGluGluSerCysGlySerSerGln 350
             :::   |||   |||   |||:::|||      :::|||:::
 891 .....GACCTGGCTCATGAGCCAGGCGAGAAAAGCTCTCTAGAGTCCAAG 935

 351 LeuHisGluAsnSerGlySerProGluThrSerArgSerLeuProAlaPr 367
        :::   :::      |||   :::   :::|||:::         ||
 936 AGTGAGGCCAGGCCCGAGTCCTCACGCCTCAAGCGCGCCTCTGCTCCCCC 985

 367 oGlnAspAsnAsp.PheLeuSerArgLysAlaGlnAspCysTyrPheMet 383
     |   |||||||||    :::|||            :::|||
 986 CTTCGATAACGACTGCAGTCTCT.........CCGAGTTGC......... 1017

 384 LysLeuHisHisCysProGlyAsnHisSerTrpAspSerThrIleSerGl 400
                 |||         |||||||||               ||
1018 ............TGT.........CACAGTTGG.........ACTCTGGG 1037

 400 ySerGlnArgAlaAlaPheCysAspHisLysThrThrProCysSerSerA 417
     ||||   |||                        :::|||   |||||||
1038 ATCTCCCAGACTCTTG...............AAGGCCCCGAAGAGCTCAG 1072

 417 laIleIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeuGlnPro 433
     ||   :::   ||||||            |||::::::      :::|||
1073 CCGAAGTT...CCTCTG............AATGCAAGCTCCCATCGTCCA 1107

 434 GlyIleAlaGln..................GlnTrpIleGlnSerLysAr 444
     ::::::|||:::                  ||||||   |||
1108 GCAGTGGCAAGAGGCTCTCGGGGGTGTCCTCAGTGGACTCAGCCTTTTCC 1157

 444 gGluAspIleValAsnGlnMetThrGluAlaCysLeuAsnGlnSerLeuA 461
         :::                  :::   |||
1158 TCCAGAG..................GATCGCTGT...............C 1174

 461 spAlaLeuLeuSerArgAspLeuIleMetLysGluAspTyrGluLeuVal 477
       :::|||||||||
1175 ACTGTCTTTTGAGC.................................... 1188

 478 SerThrLysProThrArgThrSerLysValArgGlnLeu..LeuAspThr 493
                          :::|||::::::|||     |||
1189 .....................GGGAAGCTTCAACAGGCGACCTGGGCCCC 1217

 494 ThrAspIleGlnGlyGluGluPheAlaLysValIleValGlnLysLeuLy 510
     ||||||||||||   ::::::::::::   :::|||::::::
1218 ACAGACATCCAGAAGAAGAAGCTAGTGGATGCCATCATATCAGGGGACAC 1267

 510 sAspAsnLysGlnMetGlyLeuGlnProTyrProGluIleLeuValValS 527
      ::::::   :::      |||||||||            ||||||::::
1268 CAGCAGGCTGATGAAGATCCTACAGCCCCAAGATGTGGACTTGGTTCTAG 1317

 527 erArgSerProSerLeu 532
```

```
         ::  |||  ||||||
    1318 ACAGCAGTGCCAGCCTG 1334

seq_name: /cgn1_8/gcgdata/geneseq/geneseqn/NA2001.DAT:C99763

seq_documentation_block:
ID   C99763 standard; cDNA; 1774 BP.
XX
AC   C99763;
XX
DT   08-MAR-2001  (first entry)
XX
DE   Skin cell cDNA, SEQ ID NO: 403.
XX
KW   Mouse; skin cell; cytostatic; antiinflammatory; anti-HIV;
KW   nootropic; neuroprotective; vulnerary; immunomodulatory; vaccine;
KW   keratinocyte growth stimulation; cancer; angiogenesis inhibition;
KW   inflammation; neurological disease; ss.
XX
OS   Mus sp.
XX
PN   WO200069884-A2.
XX
PD   23-NOV-2000.
XX
PF   15-MAY-2000; 2000WO-NZ00075.
XX
PR   14-MAY-1999;   99US-0312283.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Watson JD,  Strachan L,  Onrust R,  Sleeman M,  Kumble KD,  Murison JG;
XX
DR   WPI; 2001-007495/01.
DR   P-PSDB; B56062.
XX
PT   New isolated polynucleotide used in the identification of genetic
PT   disorders and encoding polypeptides used for treating inflammatory
PT   disease, cancer and neurological diseases -
XX
PS   Claim 1; Page 305-306; 352pp; English.
XX
CC   The present polynucleotide encodes a polypeptide which is expressed in
CC   mammalian skin cells. The polypeptide is useful for stimulating
CC   keratinocyte growth and motility, inhibiting the growth of cancer cells,
CC   modulating angiogenesis, inhibiting angiogenesis and vascularisation of
CC   tumours, modulating skin inflammation, stimulating the growth of
CC   epithelial cells, inhibiting the binding of human immunodeficiency virus
CC   (HIV)-1 to leukocytes, and treating inflammmatory disease, cancer and
CC   neurological diseases. The polynucleotide can be used as a marker, in
CC   the identification of genetic disorders, and for the design of
CC   oligonucleotides for examining expression patterns.
XX
SQ   Sequence 1774 BP; 406 A; 490 C; 526 G; 352 T; 0 other;


alignment_scores:
          Quality:  577.00            Length:    540
            Ratio:  1.963               Gaps:     26
 Percent Similarity:  54.444   Percent Identity: 33.889

alignment_block:
 US-09-445-223-1 x C99763   ..

 Align seg 1/1  to: C99763  from: 1  to: 1774

      10 LeuProThrIleProTyrHisLysLeuAlaAspLeuArgTyrLeuSerAr 26
         |||   |||:::         :::::::|||      :::   ::::::
      45 CTGCGCACCTTCGACGCCGGCGAATTCGCAGGCTGGGAGAAGGTGGGCTC 94

      26 gGlyAlaSerGlyThrValSerSerAlaArgHisAlaAspTrpArgValG 43
          |||:::   |||   |||   ::::::||||||:::   |||::::::
      95 GGGCGGCTTCGGGCAGGTGTACAAGGTGCGCCATGTGCACTGGAAGACGT 144
```

```
 43 lnValAlaValLys.........HisLeuHisIleHisThrProLeuLeu 56
      :::|||:::|||         ||||||:::
145 GGCTCGCGATCAAGTGCTCGCCCAGTCTGCACGTCGAC............ 182

 57 AspSerGluArgLysAspValLeuArgGluAlaGluIleLeuHisLysAl 73
    |||   ||||||   ::::::|||:::||||||:::   ::::::   ||
183 GACAGGGAACGAATGGAGCTCCTGGAGGAAGCTAAGAAGATGGAGATGGC 232

 73 aArgPheSerTyrIlePheProIleLeuGlyIleCysAsnGluProGluP 90
    |:::|||   ||||||:::|||:::   |||||||||:::||||||
233 CAAGTTCCGATACATTCTACCTGTGTACGGCATATGCCAGGAACCT.... 278

 90 heLeuGlyIleValThrGluTyrMetProAsnGlySerLeuAsnGluLeu 106
      :::|||:::|||   |||||||||   :::|||||||||::::::|||
279 ..GTCGGCTTGGTCATGGAGTACATGGAGACAGGCTCCCTGGAGAAGCTG 326

107 LeuHisArgLysThrGluTyrProAspValAlaTrpProLeuArgPheAr 123
    |||      :::               :::   |||   |||||||||||
327 CTGGCCTCAGAG...............CCATTGCCTTGGGACCTGCGCTTTCG 364

123 gIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisAsnMetThrP 140
    ||||:::||||||   |||:::|||:::|||:::||||||   |||:::|
365 CATCGTGCACGAGACAGCCGTGGGCATGAACTTCCTGCATTGCATGTCTC 414

140 roProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeuAspAsn 156
    ||||||||||||||   |||||||||      |||||||||||||||
415 CGCCACTGCTGCACCTAGACCTGAAGCCAGCGAACATCCTGCTGGATGCC 464

157 GluPheHisValLysIleAlaAspPheGlyLeuSerLysTrpArgMetMe 173
    ::::::||||||||||||:::||||||||||||:::|||   :::   ||
465 CACTACCATGTCAAGATTTCTGACTTTGGGCTGGCCAAGTGCAATGGCAT 514

173 tSerLeuSerGlnSerArgSerSerLysSerAlaProGluGly......G 188
    ||||   |||::::::   |||               :::|||      |
515 GTCCCACTCTCATGACCTCAGCATG............GATGGCCTGTTTG 552

188 lyThrIleIleTyrMetProProGluAsnTyrGluProGlyGlnLysSer 204
    ||||||||   |||:::|||||||||:::   :::      :::||||||
553 GTACAATCGCTTACCTCCCTCCAGAGCGAATTCGT......GAGAAGAGC 596

205 ArgAla...SerIleLysHisAspIleTyrSerTyrAlaValIleThrTr 220
    |||      :::   |||||||||:::||||||:::|||::::::   ||
597 CGCTTGTTTGACACCAAACATGATGTATACAGCTTCGCCATTGTGATCTG 646

220 pGluValLeuSerArgLysGlnProPheGluAspValThrAsnProLeuG 237
    |   ||||||::::::|||:::||||||   |||      |||   |||:
647 GGGTGTGCTTACACAGAAGAAGCCATTTGCAGATGAAAAGAACATCCTAC 696

237 lnIleMetTyrSerValSerGlnGlyHisArgProValIleAsnGluGlu 253
    ::||||||   :::|||   :::||||||||||||            |||
697 ACATCATGATGAAAGTGGTAAAGGGCCACCGCCCA............GAG 734

254 SerLeuProTyrAspIleProHisArg.........AlaArgMetIleSe 267
          |||         |||:::            |||   :::|||::
735 CTGCCACCCATCTGCAGACCCCGGCCGCGTGCCTGTGCCAGCCTGATAGG 784

267 rLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerPheL 284
    ::::::::::      |||      :::|||:::   ||||||
785 GATAATGCAACGGTGCTGGCATGCAGACCCACAGGTGCGGCCC....... 827

284 euLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGluIle 300
                                       ||||||:::||||||
828 ...................................ACCTTCCAAGAAATT 842

301 ThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSerVa 317
    |||
843 ACC............................................... 845

317 lSerSerAlaIleHisLeuCysAspLysLysLysMetGluLeuSerLeuA 334
     |||::::::      ||||||:::|||         |||::::::
846 .TCTGAAACAGAAGACCTTTGTGAGAAGCCTGATGAGGAGGTGAAA.... 890
```

```
 334 snIleProValAsnHisGlyProGlnGluGluSerCysGlySerSerGln 350
             :::   |||   |||   |||:::|||      :::|||:::
 891 .....GACCTGGCTCATGAGCCAGGCGAGAAAAGCTCTCTAGAGTCCAAG 935

 351 LeuHisGluAsnSerGlySerProGluThrSerArgSerLeuProAlaPr 367
        :::   :::      |||   :::   :::|||:::         ||
 936 AGTGAGGCCAGGCCCGAGTCCTCACGCCTCAAGCGCGCCTCTGCTCCCCC 985

 367 oGlnAspAsnAsp.PheLeuSerArgLysAlaGlnAspCysTyrPheMet 383
     |   |||||||||    :::|||            :::|||
 986 CTTCGATAACGACTGCAGTCTCT.........CCGAGTTGC......... 1017

 384 LysLeuHisHisCysProGlyAsnHisSerTrpAspSerThrIleSerGl 400
                 |||         |||||||||                ||
1018 ............TGT.........CACAGTTGG..........ACTCTGGG 1037

 400 ySerGlnArgAlaAlaPheCysAspHisLysThrThrProCysSerSerA 417
     ||||   |||                        :::|||   |||||||
1038 ATCTCCCAGACTCTTG...............AAGGCCCCGAAGAGCTCAG 1072

 417 laIleIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeuGlnPro 433
     ||   :::   ||||||            |||::::::      :::|||
1073 CCGAAGTT...CCTCTG............AATGCAAGCTCCCATCGTCCA 1107

 434 GlyIleAlaGln..................GlnTrpIleGlnSerLysAr 444
     ::::::|||:::                  ||||||   |||
1108 GCAGTGGCAAGAGGCTCTCGGGGGTGTCCTCAGTGGACTCAGCCTTTTCC 1157

 444 gGluAspIleValAsnGlnMetThrGluAlaCysLeuAsnGlnSerLeuA 461
         :::                  :::   |||
1158 TCCAGAG..................GATCGCTGT...............C 1174

 461 spAlaLeuLeuSerArgAspLeuIleMetLysGluAspTyrGluLeuVal 477
       :::|||||||||
1175 ACTGTCTTTTGAGC.................................... 1188

 478 SerThrLysProThrArgThrSerLysValArgGlnLeu..LeuAspThr 493
                          :::|||::::::|||     |||
1189 .....................GGGAAGCTTCAACAGGCGACCTGGGCCCC 1217

 494 ThrAspIleGlnGlyGluGluPheAlaLysValIleValGlnLysLeuLy 510
     ||||||||||||   ::::::::::::   :::|||::::::
1218 ACAGACATCCAGAAGAAGAAGCTAGTGGATGCCATCATATCAGGGGACAC 1267

 510 sAspAsnLysGlnMetGlyLeuGlnProTyrProGluIleLeuValValS 527
      ::::::   :::      |||||||||            ||||||::::
1268 CAGCAGGCTGATGAAGATCCTACAGCCCCAAGATGTGGACTTGGTTCTAG 1317

 527 erArgSerProSerLeu 532
     ::   |||   ||||||
1318 ACAGCAGTGCCAGCCTG 1334

seq_name: /cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:Z58584

seq_documentation_block:
ID   Z58584 standard; cDNA; 2294 BP.
XX
AC   Z58584;
XX
DT   05-JUN-2000  (first entry)
XX
DE   Mouse protein kinase Fel cDNA.
XX
KW   Fel; protein kinase; mouse; signal transduction; ss.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             3..2294
FT                   /*tag=  a
XX
PN   WO200008178-A2.
XX
PD   17-FEB-2000.
XX
PF   04-AUG-1999;   99WO-US17577.
XX
PR   04-AUG-1998;   98US-0095269.
PR   11-SEP-1998;   98US-0099973.
XX
PA   (IMMV ) IMMUNEX CORP.
XX
PI   Virca GD,  Bird TA,  Anderson DM,  Marken JS;
XX
DR   WPI; 2000-205722/18.
DR   P-PSDB; Y79154.
XX
PT   Novel murine polynucleotides encoding kinase polypeptides, used as
PT   probes to identify nucleic acids encoding proteins having kinase
PT   activity -
XX
PS   Claim 1; Fig 3; 93pp; English.
XX
CC   The present sequence is that of Genesis clone 971025TRAM004820HT
CC   (Fel) DNA.  The translation of the clone (frame 3) is provided in
CC   Y79154.  The polypeptide has kinase activity.  The invention
CC   relates to purified murine polypeptides (see Y79152-56) that
CC   have kinase function and isolated nucleic acids encoding them (see
CC   Z58582-86).  The nucleic acids can be used to express the
CC   polypeptides, as probes to identify nucleic acids encoding proteins
CC   having kinase activity, and in assays to identify chromosomes; map
CC   murine genes, and study tumours.  Vectors comprising the novel
CC   nucleic acids and host cells (bacterial, yeast, plant, insect or
CC   animal) transfected or transduced with the vectors are claimed.
CC   The kinase polypeptides and their fragments are used as mol.wt. and
CC   isoelectric focusing markers, and as controls for peptide
CC   fragmentation.  They also have a number of therapeutic uses, as
CC   kinases play a central role in cellular signal transduction.
XX
SQ   Sequence 2294 BP; 526 A; 639 C; 666 G; 463 T; 0 other;

alignment_scores:
          Quality:  577.00             Length:    523
            Ratio:  2.025                Gaps:     26
 Percent Similarity:  54.493   Percent Identity: 34.608

alignment_block:
 US-09-445-223-1 x Z58584 .  ..

 Align seg 1/1  to: Z58584  from: 1  to: 2294

  27 GlyAlaSerGlyThrValSerSerAlaArgHisAlaAspTrpArgValGl 43
     |||:::   |||   |||   ::::::|||||||:::  |||::::::
  24 GGCGGCTTCGGGCAGGTGTACAAGGTGCGCCATGTGCACTGGAAGACGTG 73

  43 nValAlaValLys.........HisLeuHisIleHisThrProLeuLeuA 57
      :::|||:::|||            |||||||:::              |
  74 GCTCGCGATCAAGTGCTCGCCCAGTCTGCACGTCGAC............G 111

  57 spSerGluArgLysAspValLeuArgGluAlaGluIleLeuHisLysAla 73
     ||   ||||||   ::::::|||:::|||||||:::   ::::::  |||
 112 ACAGGGAACGAATGGAGCTCCTGGAGGAAGCTAAGAAGATGGAGATGGCC 161

  74 ArgPheSerTyrIlePheProIleLeuGlyIleCysAsnGluProGluPh 90
     :::|||   ||||||:::|||:::   |||||||||:::||||||
 162 AAGTTCCGATACATTCTACCTGTGTACGGCATATGCCAGGAACCT..... 206

  90 eLeuGlyIleValThrGluTyrMetProAsnGlySerLeuAsnGluLeuL 107
      :::|||:::|||   |||||||||   :::|||||||||::::::||||
 207 .GTCGGCTTGGTCATGGAGTACATGGAGACAGGCTCCCTGGAGAAGCTGC 255

 107 euHisArgLysThrGluTyrProAspValAlaTrpProLeuArgPheArg 123
     ||      :::               :::   |||   ||||||||||||
 256 TGGCCTCAGAG............CCATTGCCTTGGGACCTGCGCTTTCGC 293
```

```
124 IleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisAsnMetThrPr 140
    |||:::||||||   |||:::|||:::|||:::||||||   |||:::||
294 ATCGTGCACGAGACAGCCGTGGGCATGAACTTCCTGCATTGCATGTCTCC 343

140 oProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeuAspAsnG 157
    |||||||||||||   |||||||||      |||||||||||||||   :
344 GCCACTGCTGCACCTAGACCTGAAGCCAGCGAACATCCTGCTGGATGCCC 393

157 luPheHisValLysIleAlaAspPheGlyLeuSerLysTrpArgMetMet 173
    :::::||||||||||||:::||||||||||||:::|||   :::   |||
394 ACTACCATGTCAAGATTTCTGACTTTGGGCTGGCCAAGTGCAATGGCATG 443

174 SerLeuSerGlnSerArgSerSerLysSerAlaProGluGly......Gl 188
    |||   |||::::::   |||               :::|||      ||
444 TCCCACTCTCATGACCTCAGCATG............GATGGCCTGTTTGG 481

188 yThrIleIleTyrMetProProGluAsnTyrGluProGlyGlnLysSerA 205
    |||||||   |||:::|||||||||:::   :::      :::|||||||
482 TACAATCGCTTACCTCCCTCCAGAGCGAATTCGT......GAGAAGAGCC 525

205 rgAla...SerIleLysHisAspIleTyrSerTyrAlaValIleThrTrp 220
    ||      :::   |||||||||:::||||||:::|||::::::   |||
526 GCTTGTTTGACACCAAACATGATGTATACAGCTTCGCCATTGTGATCTGG 575

221 GluValLeuSerArgLysGlnProPheGluAspValThrAsnProLeuGl 237
       ||||||::::::|||:::||||||   |||      |||   |||::
576 GGTGTGCTTACACAGAAGAAGCCATTTGCAGATGAAAAGAACATCCTACA 625

237 nIleMetTyrSerValSerGlnGlyHisArgProValIleAsnGluGluS 254
    :||||||   :::|||   :::||||||||||||            |||
626 CATCATGATGAAAGTGGTAAAGGGCCACCGCCCA............GAGC 663

254 erLeuProTyrAspIleProHisArg.........AlaArgMetIleSer 267
         |||         |||:::            |||   :::|||:::
664 TGCCACCCATCTGCAGACCCCGGCCGCGTGCCTGTGCCAGCCTGATAGGG 713

268 LeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerPheLe 284
    |||::::::      |||      :::|||:::   ||||||
714 CTCATGCAACGGTGCTGGCATGCAGACCCACAGGTGCGGCCC........ 755

284 uLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGluIleT 301
                                      ||||||:::|||||||
756 ..................................ACCTTCCAAGAAATTA 771

301 hrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSerVal 317
    ||
772 CC................................................ 773

318 SerSerAlaIleHisLeuCysAspLysLysLysMetGluLeuSerLeuAs 334
    |||::::::      ||||||:::|||         |||::::::
774 TCTGAAACAGAAGACCTTTGTGAGAAGCCTGATGAGGAGGTGAAA..... 818

334 nIleProValAsnHisGlyProGlnGluGluSerCysGlySerSerGlnL 351
           :::   |||   |||   |||:::|||      :::|||:::
819 ....GACCTGGCTCATGAGCCAGGCGAGAAAAGCTCTCTAGAGTCCAAGA 864

351 euHisGluAsnSerGlySerProGluThrSerArgSerLeuProAlaPro 367
      :::   :::      |||   :::   :::|||:::         |||
865 GTGAGGCCAGGCCCGAGTCCTCACGCCTCAAGCGCGCCTCTGCTCCCCCC 914

368 GlnAspAsnAsp.PheLeuSerArgLysAlaGlnAspCysTyrPheMetL 384
       |||||||||    :::|||            :::|||
915 TTCGATAACGACTGCAGTCTCT.........CCGAGTTGC.......... 945

384 ysLeuHisHisCysProGlyAsnHisSerTrpAspSerThrIleSerGly 400
               |||         |||||||||               |||
946 ...........TGT.........CACAGTTGG.........ACTCTGGGA 966

401 SerGlnArgAlaAlaPheCysAspHisLysThrThrProCysSerSerAl 417
    |||   |||                        :::|||   ||||||||
967 TCTCCCAGACTCTTG...............AAGGCCCCGAAGAGCTCAGC 1001

417 aIleIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeuGlnProG 434
     |   :::   ||||||            |||::::::      :::|||:
1002 CGAAGTT...CCTCTG............AATGCAAGCTCCCATCGTCCAG 1036

 434 lyIleAlaGln...................GlnTrpIleGlnSerLysArg 444
     :::::|||:::                  ||||||   |||
1037 CAGTGGCAAGAGGCTCTCGGGGGTGTCCTCAGTGGACTCAGCCTTTTCCT 1086

 445 GluAspIleValAsnGlnMetThrGluAlaCysLeuAsnGlnSerLeuAs 461
        :::                  :::   |||
1087 CCAGAG..................GATCGCTGT...............CA 1103

 461 pAlaLeuLeuSerArgAspLeuIleMetLysGluAspTyrGluLeuValS 478
      :::|||||||||
1104 CTGTCTTTTGAGC..................................... 1116

 478 erThrLysProThrArgThrSerLysValArgGlnLeu..LeuAspThrT 494
                         :::|||::::::|||     |||      |
1117 ....................GGGAAGCTTCAACAGGCGACCTGGGCCCCA 1146

 494 hrAspIleGlnGlyGluGluPheAlaLysValIleValGlnLysLeuLys 510
     |||||||||||   ::::::::::::   :::|||::::::
1147 CAGACATCCAGAAGAAGAAGCTAGTGGATGCCATCATATCAGGGGACACC 1196

 511 AspAsnLysGlnMetGlyLeuGlnProTyrProGluIleLeuValValSe 527
     ::::::   :::      |||||||||            ||||||:::::
1197 AGCAGGCTGATGAAGATCCTACAGCCCCAAGATGTGGACTTGGTTCTAGA 1246

 527 rArgSerProSerLeu 532
     :   |||   ||||||
1247 CAGCAGTGCCAGCCTG 1262

seq_name: /cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:Z61671

seq_documentation_block:
ID   Z61671 standard; cDNA; 1888 BP.
XX
AC   Z61671;
XX
DT   27-MAR-2000  (first entry)
XX
DE   cDNA encoding murine RIP protein kinase homologue, SEQ ID NO:66.
XX
KW   Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
KW   embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
KW   secreted; transmembrane; inflammation; cancer; neurological disease;
KW   angiogenesis; tumour vascularisation; growth disorder;
KW   developmental disorder; skin wound; hair follicle disorder;
KW   anti-inflammatory; cytostatic; neuroprotective; vulnery; ss.
XX
OS   Mus sp.
XX
PN   WO9955865-A1.
XX
PD   04-NOV-1999.
XX
PF   29-APR-1999;   99WO-NZ00051.
XX
PR   29-APR-1998;   98US-0069726.
PR   09-NOV-1998;   98US-0188930.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Strachan L,  Sleeman M,  Watson JD,  Onrust R,  Kumble A,  Murison JG;
XX
DR   WPI; 2000-072177/06.
DR   P-PSDB; Y76007.
XX
PT   Novel polynucleotides useful for the treatment of various conditions
PT   including wounds and cancer -
XX
PS   Claim 1; Page 85-86; 235pp; English.
XX
CC   The invention relates to novel nucleic acid sequences derived from rat
```

```
CC   dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
CC   and mouse embryonic skin, keratinocyte stem cells and transit amplifying
CC   cells. Polypeptides of the invention may be used to treat inflammation,
CC   cancer and neurological diseases. The proteins may be used to stimulate
CC   the growth and motility of keratinocytes, to inhibit the growth of
CC   cancer cells, to modulate angiogenesis and tumour vascularisation, to
CC   modulate skin inflammation, to modulate epithelial cell growth and to
CC   inhibit binding of HIV-1 to leukocytes. The invention may also be used to
CC   treat growth and developmental defects, skin wounds and hair follicle
CC   disorders. Sequences Z61606-Z61832 represent cDNA sequences derived from
CC   several mouse, rat or human skin cell types. Sequences Z61606-Z61649,
CC   Z61725-Z61765, Z61802-Z61811 and Z61826 encode proteins with an
CC   N-terminal signal sequence, indicating that the proteins are secreted.
CC   Sequences Z61650-Z61668, Z61766-Z61780, Z61812-Z61817 and Z61827-Z61829
CC   encode proteins with one or more putative transmembrane domains.
XX
SQ   Sequence 1888 BP; 433 A; 518 C; 554 G; 380 T; 3 other;


alignment_scores:
           Quality:  548.00              Length:    541
             Ratio:  1.845                 Gaps:     27
 Percent Similarity:  54.898   Percent Identity: 33.272

alignment_block:
 US-09-445-223-1 x Z61671   ..

 Align seg 1/1  to: Z61671  from: 1  to: 1888

  10 LeuProThrIleProTyrHisLysLeuAlaAspLeuArgTyrLeuSerAr 26
     |||   |||:::         ::::::|||       :::   ::::::
  43 CTGCGCACCTTCGACGCCGGCGAATTCGCAGGCTGGGAGAAGGTGGGCTC 92

  26 gGlyAlaSerGlyThrValSerSerAlaArgHisAlaAspTrpArgValG 43
      |||:::   |||   |||   ::::::||||||:::   |||::::::
  93 GGGCGGCTTCGGGCAGGTGTACAAGGTGCGCCATGTGCACTGGAAGACGT 142

  43 lnValAlaValLys.........HisLeuHisIleHisThrProLeuLeu 56
       :::|||:::|||            ||||||:::
 143 GGCTCGCGATCAAGTGCTCGCCCAGTCTGCACGTCGAC............ 180

  57 AspSerGluArgLysAspValLeuArgGluAlaGluIleLeuHisLysAl 73
     |||   ||||||   ::::::|||:::||||||:::   ::::::   ||
 181 GACAGGGAACGAATGGAGCTCCTGGAGGAAGCTAAGAAGATGGAGATGGC 230

  73 aArgPheSerTyrIlePheProIleLeuGlyIleCysAsnGluProGluP 90
     |:::|||   ||||||:::|||:::   |||||||||:::||||||
 231 CAAGTTCCGATACATTCTACCTGTGTACGGCATATGCCAGGAACCT.... 276

  90 heLeuGlyIleValThrGluTyrMetProAsnGlySerLeuAsnGluLeu 106
       :::|||:::|||   |||||||||   :::|||||||||::::::|||
 277 ..GTCGGCTTGGTCATGGAGTACATGGAGACAGGCTCCCTGGAGAAGCTG 324

 107 LeuHisArgLysThrGluTyrProAspValAlaTrpProLeuArgPheAr 123
     |||      :::               :::   |||   |||||||||||
 325 CTGGCCTCAGAG..............CCATTGCCTTGGGACCTGCGCTTTCG 362

 123 gIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisAsnMetThrP 140
     ||||:::||||||   |||:::|||:::|||:::||||||   |||:::|
 363 CATCGTGCACGAGACAGCCGTGGGCATGAACTTCCTGCATTGCATGTCTC 412

 140 roProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeuAspAsn 156
     ||||||||||||||   |||||||||      |||||||||||||||
 413 CGCCACTGCTGCACCTAGACCTGAAGCCAGCGAACATCTTGCTGGATGCC 462

 157 GluPheHisVal...LysIleAlaAspPheGlyLeuSerLysTrpArgMe 172
     ::::::::::::   ::::::   ||||||||||||:::|||   :::
 463 CACTACCAAATGTCAAGATTTCTTGACTTTGGGCTGGCCAAGTGCAATGG 512

 172 tMetSerLeuSerGlnSerArgSerSerLysSerAlaProGluGly.... 187
      ||||||   |||::::::   |||               :::|||
 513 CATGTCCCACTCTCATGACCTCAGCATG...........GATGGCCTGT 550

 188 ..GlyThrIleIleTyrMetProProGluAsnTyrGluProGlyGlnLys 203
       |||||||||   |||:::|||||||||:::   :::      :::|||
 551 TTGGTACAATCGGCTACCTCCCTCCAGAGCGAATTCGT......GAGAAG 594

 204 SerArgAla...SerIleLysHisAspIleTyrSerTyrAlaValIleTh 219
     ||||||      :::   |||||||||:::||||||:::|||::::::
 595 AGCCGCTTGTTTGACACCAAACATGATGTATACAGCTTCGCCATTGTGAT 644

 219 rTrpGluValLeuSerArgLysGlnProPheGluAspValThrAsnProL 236
      |||   ||||||::::::::::::||||||   |||      |||   |
 645 CTGGGGTGTGCTTACACAGAATAATCCATTTGCAGATGAAAAGAACATCC 694

 236 euGlnIleMetTyrSerValSerGlnGlyHisArgProValIleAsnGlu 252
     ||:::||||||   :::|||   :::||||||||||||
 695 TACACATCATGATGAAAGTGGTAAAGGGCCACCGCCCA............ 732

 253 GluSerLeuProTyrAspIleProHisArg.........AlaArgMetIl 266
     |||      |||         |||:::            |||   :::||
 733 GAGCTGCCACCCATCTGCAGACCCCGGCCGCGTGCCTGTGCCAGCCTGAT 782

 266 eSerLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerP 283
     |:::|||::::::      |||      :::|||:::   ||||||
 783 AGGGCTCATGCAACGGTGCTGGCATGCAGACCCACAGGTGCGGCCC.... 828

 283 heLeuLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGlu 299
                                           ||||||:::|||
 829 ......................................ACCTTCCAAGAA 840

 300 IleThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSe 316
     ||||||
 841 ATTACC............................................ 846

 316 rValSerSerAlaIleHisLeuCysAspLysLysLysMetGluLeuSerL 333
         |||::::::      ||||||:::|||         |||::::::
 847 ....TCTGAAACAGAAGACCTTTGTGAGAAGCCTGATGAGGAGGTGAAA. 891

 333 euAsnIleProValAsnHisGlyProGlnGluGluSerCysGlySerSer 349
                :::   |||   |||   |||:::|||      :::|||
 892 ........GACCTGGCTCATGAGCCAGGCGAGAAAAGCTCTCTAGAGTCC 933

 350 GlnLeuHisGluAsnSerGlySerProGluThrSerArgSerLeuProAl 366
     :::   :::   :::      |||   :::   :::|||:::
 934 AAGAGTGAGGCCAGGCCCGAGTCCTCACGCCTCAAGCGCGCCTCTGCTCC 983

 366 aProGlnAspAsnAsp.PheLeuSerArgLysAlaGlnAspCysTyrPhe 382
      |||   |||||||||    :::|||            :::|||
 984 CCCCTTCGATAACGACTGCAGTCTCT.........CCGAGTTGC...... 1018

 383 MetLysLeuHisHisCysProGlyAsnHisSerTrpAspSerThrIleSe 399
                    |||         |||||||||      |||:::::
1019 ...............TGT.........CACAGTTGG......ACTCTGGG 1038

 399 rGlySerGlnArgAlaAlaPheCysAspHisLysThrThrProCysSerS 416
     ::::||||||               |||         :::|||   ||||
1039 ATCTTCCCAA...............GACTCTTGAAAGGCCCCGAAGAGCT 1073

 416 erAlaIleIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeuGln 432
     |||||   :::   ||||||            |||::::::      :::
1074 CAGCCGAAGTT...CCTCTG............AATGCAAGCTCCCATCGT 1108

 433 ProGlyIleAlaGln..................GlnTrpIleGlnSerLy 443
     |||::::::|||:::                  ||||||   |||
1109 CCAGCAGTGGCAAGAGGCTCTCGGGGGTGTCCTCAGTGGACTCAGCCTTT 1158

 443 sArgGluAspIleValAsnGlnMetThrGluAlaCysLeuAsnGlnSerL 460
            :::                  :::   |||
1159 TCCTCCAGAG..................GATCGCTGT............. 1177

 460 euAspAlaLeuLeuSerArgAspLeuIleMetLysGluAspTyrGluLeu 476
          :::|||||||||
1178 ..CACTGTCTTTTGAGC................................. 1192

 477 ValSerThrLysProThrArgThrSerLysValArgGlnLeu..LeuAsp 492
```

```
                              :::|||::::::|||     |||
 1193 ........................GGGAAGCTTCAACAGGCGACCTGGGC 1218

  493 ThrThrAspIleGlnGlyGluGluPheAlaLysValIleValGlnLysLe 509
         ||||||||||||   ::::::::::::   :::|||::::::
 1219 CCCACAGACATCCAGAAGAAGAAGCTAGTGGATGCCATCATATCAGGGGA 1268

  509 uLysAspAsnLysGlnMetGlyLeuGlnProTyrProGluIleLeuValV 526
          ::::::   :::      |||||||||            ||||||:
 1269 CACCAGCAGGCTGATGAAGATCCTACAGCCCCAAGATGTGGACTTGGTTC 1318

  526 alSerArgSerProSerLeu 532
      :::::   |||   ||||||
 1319 TAGACAGCAGTGCCAGCCTG 1338

seq_name: /cgn1_8/gcgdata/geneseq/geneseqn/NA2001.DAT:C99604

seq_documentation_block:
ID   C99604 standard; cDNA; 1888 BP.
XX
AC   C99604;
XX
DT   08-MAR-2001  (first entry)
XX
DE   Skin cell cDNA, SEQ ID NO: 66.
XX
KW   Mouse; skin cell; cytostatic; antiinflammatory; anti-HIV;
KW   nootropic; neuroprotective; vulnerary; immunomodulatory; vaccine;
KW   keratinocyte growth stimulation; cancer; angiogenesis inhibition;
KW   inflammation; neurological disease; ss.
XX
OS   Mus sp.
XX
PN   WO200069884-A2.
XX
PD   23-NOV-2000.
XX
PF   15-MAY-2000; 2000WO-NZ00075.
XX
PR   14-MAY-1999;   99US-0312283.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Watson JD,  Strachan L,  Onrust R,  Sleeman M,  Kumble KD,  Murison JG;
XX
DR   WPI; 2001-007495/01.
DR   P-PSDB; B55946.
XX
PT   New isolated polynucleotide used in the identification of genetic
PT   disorders and encoding polypeptides used for treating inflammatory
PT   disease, cancer and neurological diseases -
XX
PS   Claim 1; Page 109-110; 352pp; English.
XX
CC   The present polynucleotide encodes a polypeptide which is expressed in
CC   mammalian skin cells. The polypeptide is useful for stimulating
CC   keratinocyte growth and motility, inhibiting the growth of cancer cells,
CC   modulating angiogenesis, inhibiting angiogenesis and vascularisation of
CC   tumours, modulating skin inflammation, stimulating the growth of
CC   epithelial cells, inhibiting the binding of human immunodeficiency virus
CC   (HIV)-1 to leukocytes, and treating inflammmatory disease, cancer and
CC   neurological diseases. The polynucleotide can be used as a marker, in
CC   the identification of genetic disorders, and for the design of
CC   oligonucleotides for examining expression patterns.
XX
SQ   Sequence 1888 BP; 433 A; 518 C; 554 G; 380 T; 3 other;

alignment_scores:
           Quality:  548.00            Length:    541
             Ratio:  1.845               Gaps:     27
 Percent Similarity:  54.898   Percent Identity: 33.272
```

```
alignment_block:
 US-09-445-223-1 x C99604   ..

 Align seg 1/1  to: C99604  from: 1  to: 1888

   10 LeuProThrIleProTyrHisLysLeuAlaAspLeuArgTyrLeuSerAr 26
      |||   |||:::         ::::::|||       :::   ::::::
   43 CTGCGCACCTTCGACGCCGGCGAATTCGCAGGCTGGGAGAAGGTGGGCTC 92

   26 gGlyAlaSerGlyThrValSerSerAlaArgHisAlaAspTrpArgValG 43
       |||:::   |||   |||   ::::::|||||||:::   |||::::::
   93 GGGCGGCTTCGGGCAGGTGTACAAGGTGCGCCATGTGCACTGGAAGACGT 142

   43 lnValAlaValLys.........HisLeuHisIleHisThrProLeuLeu 56
        :::|||:::|||             ||||||:::
  143 GGCTCGCGATCAAGTGCTCGCCCAGTCTGCACGTCGAC............ 180

   57 AspSerGluArgLysAspValLeuArgGluAlaGluIleLeuHisLysAl 73
      |||   ||||||    ::::::|||:::||||||:::   ::::::   ||
  181 GACAGGGAACGAATGGAGCTCCTGGAGGAAGCTAAGAAGATGGAGATGGC 230

   73 aArgPheSerTyrIlePheProIleLeuGlyIleCysAsnGluProGluP 90
      |:::|||   ||||||:::|||:::    |||||||||:::||||||
  231 CAAGTTCCGATACATTCTACCTGTGTACGGCATATGCCAGGAACCT.... 276

   90 heLeuGlyIleValThrGluTyrMetProAsnGlySerLeuAsnGluLeu 106
        :::|||:::|||   |||||||||    :::|||||||||::::::|||
  277 ..GTCGGCTTGGTCATGGAGTACATGGAGACAGGCTCCCTGGAGAAGCTG 324

  107 LeuHisArgLysThrGluTyrProAspValAlaTrpProLeuArgPheAr 123
      |||      :::                :::   |||   |||||||||||
  325 CTGGCCTCAGAG.............CCATTGCCTTGGGACCTGCGCTTTCG 362

  123 gIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisAsnMetThrP 140
      ||||:::||||||   |||:::|||:::|||:::|||||||   |||:::|
  363 CATCGTGCACGAGACAGCCGTGGGCATGAACTTCCTGCATTGCATGTCTC 412

  140 roProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeuAspAsn 156
      ||||||||||||||   |||||||||       |||||||||||||||
  413 CGCCACTGCTGCACCTAGACCTGAAGCCAGCGAACATCTTGCTGGATGCC 462

  157 GluPheHisVal...LysIleAlaAspPheGlyLeuSerLysTrpArgMe 172
      ::::::::::::   ::::::    ||||||||||||:::|||   :::
  463 CACTACCAAATGTCAAGATTTCTTGACTTTGGGCTGGCCAAGTGCAATGG 512

  172 tMetSerLeuSerGlnSerArgSerSerLysSerAlaProGluGly.... 187
       ||||||   |||::::::   |||                :::|||
  513 CATGTCCCACTCTCATGACCTCAGCATG............GATGGCCTGT 550

  188 ..GlyThrIleIleTyrMetProProGluAsnTyrGluProGlyGlnLys 203
        |||||||||   |||:::|||||||||:::    :::      :::|||
  551 TTGGTACAATCGGCTACCTCCCTCCAGAGCGAATTCGT......GAGAAG 594

  204 SerArgAla...SerIleLysHisAspIleTyrSerTyrAlaValIleTh 219
      ||||||      :::   |||||||||:::||||||:::|||::::::
  595 AGCCGCTTGTTTGACACCAAACATGATGTATACAGCTTCGCCATTGTGAT 644

  219 rTrpGluValLeuSerArgLysGlnProPheGluAspValThrAsnProL 236
       |||   ||||||::::::::::::||||||    |||      |||   |
  645 CTGGGGTGTGCTTACACAGAATAATCCATTTGCAGATGAAAAGAACATCC 694

  236 euGlnIleMetTyrSerValSerGlnGlyHisArgProValIleAsnGlu 252
      ||:::||||||   :::|||   :::||||||||||||
  695 TACACATCATGATGAAAGTGGTAAAGGGCCACCGCCCA............ 732

  253 GluSerLeuProTyrAspIleProHisArg.........AlaArgMetIl 266
      |||      |||         |||:::             |||   :::||
  733 GAGCTGCCACCCATCTGCAGACCCCGGCCGCGTGCCTGTGCCAGCCTGAT 782

  266 eSerLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerP 283
      |:::|||::::::      |||       :::|||:::    ||||||
  783 AGGGCTCATGCAACGGTGCTGGCATGCAGACCCACAGGTGCGGCCC.... 828
```

```
    283 heLeuLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGlu 299
                                                ||||||:::|||
    829 ......................................ACCTTCCAAGAA 840

    300 IleThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSe 316
        ||||||
    841 ATTACC............................................ 846

    316 rValSerSerAlaIleHisLeuCysAspLysLysLysMetGluLeuSerL 333
            |||::::::      ||||||:::|||         |||::::::
    847 ....TCTGAAACAGAAGACCTTTGTGAGAAGCCTGATGAGGAGGTGAAA. 891

    333 euAsnIleProValAsnHisGlyProGlnGluGluSerCysGlySerSer 349
                   :::   |||   |||   |||:::|||      :::|||
    892 ........GACCTGGCTCATGAGCCAGGCGAGAAAAGCTCTCTAGAGTCC 933

    350 GlnLeuHisGluAsnSerGlySerProGluThrSerArgSerLeuProAl 366
        :::   :::   :::      |||   :::   :::|||:::
    934 AAGAGTGAGGCCAGGCCCGAGTCCTCACGCCTCAAGCGCGCCTCTGCTCC 983

    366 aProGlnAspAsnAsp.PheLeuSerArgLysAlaGlnAspCysTyrPhe 382
         |||   |||||||||    :::|||            :::|||
    984 CCCCTTCGATAACGACTGCAGTCTCT.........CCGAGTTGC...... 1018

    383 MetLysLeuHisHisCysProGlyAsnHisSerTrpAspSerThrIleSe 399
                       |||         |||||||||      |||:::::
   1019 ...............TGT.........CACAGTTGG......ACTCTGGG 1038

    399 rGlySerGlnArgAlaAlaPheCysAspHisLysThrThrProCysSerS 416
        ::::||||||               |||         :::|||   ||||
   1039 ATCTTCCCAA...............GACTCTTGAAAGGCCCCGAAGAGCT 1073

    416 erAlaIleIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeuGln 432
        |||||   :::   ||||||            |||::::::      :::
   1074 CAGCCGAAGTT...CCTCTG............AATGCAAGCTCCCATCGT 1108

    433 ProGlyIleAlaGln.................GlnTrpIleGlnSerLy 443
        |||::::::|||:::                  ||||||   |||
   1109 CCAGCAGTGGCAAGAGGCTCTCGGGGGTGTCCTCAGTGGACTCAGCCTTT 1158

    443 sArgGluAspIleValAsnGlnMetThrGluAlaCysLeuAsnGlnSerL 460
               :::                  :::   |||
   1159 TCCTCCAGAG..................GATCGCTGT............. 1177

    460 euAspAlaLeuLeuSerArgAspLeuIleMetLysGluAspTyrGluLeu 476
             :::|||||||||
   1178 ..CACTGTCTTTTGAGC................................. 1192

    477 ValSerThrLysProThrArgThrSerLysValArgGlnLeu..LeuAsp 492
                                :::|||::::::|||     |||
   1193 ........................GGGAAGCTTCAACAGGCGACCTGGGC 1218

    493 ThrThrAspIleGlnGlyGluGluPheAlaLysValIleValGlnLysLe 509
           ||||||||||||   ::::::::::::   :::|||::::::
   1219 CCCACAGACATCCAGAAGAAGAAGCTAGTGGATGCCATCATATCAGGGGA 1268

    509 uLysAspAsnLysGlnMetGlyLeuGlnProTyrProGluIleLeuValV 526
            ::::::   :::      |||||||||            ||||||:
   1269 CACCAGCAGGCTGATGAAGATCCTACAGCCCCAAGATGTGGACTTGGTTC 1318

    526 alSerArgSerProSerLeu 532
        :::::   |||   ||||||
   1319 TAGACAGCAGTGCCAGCCTG 1338

seq_name: /cgn1_8/gcgdata/geneseq/geneseqn/NA1997.DAT:T43752

seq_documentation_block:
ID   T43752 standard; cDNA; 2268 BP.
XX
AC   T43752;
XX
DT   13-FEB-1997  (first entry)
XX
```

```
DE   Mouse receptor interacting protein gene.
XX
KW   Receptor interacting protein; RIP; cell death; apoptosis;
KW   Fas; APO-1; signal transduction; cancer; gene therapy; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             52..2022
FT                   /*tag=  a
XX
PN   WO9636730-A1.
XX
PD   21-NOV-1996.
XX
PF   18-APR-1996;   96WO-US05386.
XX
PR   18-MAY-1995;   95US-0444005.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
PA   (HARD ) HARVARD COLLEGE.
XX
PI   Kim E,  Leder P,  Lee T,  Seed B,  Stranger BZ;
XX
DR   WPI; 1997-012100/01.
DR   P-PSDB; W04627.
XX
PT   Receptor interacting protein having death and kinase domain - useful
PT   to control diseases that involve abnormal apoptosis, and for
PT   diagnosis and drug screening
XX
PS   Disclosure; Page 38-39; 64pp; English.
XX
CC   A cDNA clone (T43752) codes for mouse receptor interacting protein
CC   (RIP) (W04627), a protein which contains a C-terminal death domain
CC   through which RIP interacts with the Fas/APO-1 intracellular domain
CC   (ICD).  RIP overexpression leads to cell death.  The RIP cDNA was
CC   identified from a mouse thymus cDNA library by screening with the
CC   human RIP coding sequence (see also T43753) and sequence analysis
CC   of overlapping clones.  It can be used as a probe and to produce
CC   recombinant RIP in host cells.
XX
SQ   Sequence 2268 BP; 719 A; 491 C; 576 G; 482 T; 0 other;


alignment_scores:
           Quality:  397.00             Length:    609
             Ratio:  1.272                Gaps:     25
 Percent Similarity:  51.232   Percent Identity: 25.452

alignment_block:
 US-09-445-223-1 x T43752   ..

 Align seg 1/1  to: T43752  from: 1  to: 2268

      24 LeuSerArgGlyAlaSerGlyThrValSerSerAlaArgHisAlaAspTr 40
         |||:::   |||:::   |||   ||||||   :::   |||   :::
     118 CTAGACAGCGGAGGCTTCGGGAAGGTGTCCTTGTGTTACCACAGAAGCCA 167

      40 pArgValGlnValAlaValLysHisLeuHisIleHisThrProLeuLeuA 57
                   :::      |||         ::::::|||
     168 TGGATTTGTCATCCTGAAAAAA.........GTATACACAGGGCCCAACC 208

      57 spSerGluArgLysAspVal...LeuArgGluAlaGluIleLeuHisLys 72
           :::|||   ::::::|||   |||:::|||::::::::::::|||:::
     209 GCGCTGAGTACAATGAGGTTCTCTTGGAAGAGGGGAAGATGATGCACAGA 258

      73 AlaArgPheSerTyrIlePheProIleLeuGlyIleCysAsnGluProGl 89
            |||   |||   :::      :::|||||||||      |||
     259 CTGAGACACAGTCGAGTGGTGAAGCTACTGGGCATCATCATAGAAGAAGG 308

      89 uPheLeuGlyIleValThrGluTyrMetProAsnGlySerLeuAsnGluL 106
                ::::::|||   |||||||||   :::|||:::|||   ::::
```

```
 309 GAACTATTCGCTGGTGATGGAGTACATGGAGAAGGGCAACCTGATGCACG 358

 106 euLeuHisArgLysThrGluTyrProAspValAlaTrpProLeuArgPhe 122
     ::|||      ||||||:::      ||||||         |||:::
 359 TGCTA......AAGACCCAGATA...GATGTCCCACTTTCATTGAAAGGA 399

 123 ArgIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisAsnMetTh 139
     ||||||:::   |||         |||:::   |||||||||:::
 400 AGGATAATCGTGGAGGCCATAGAAGGCATGTGCTACTTACATGAC..... 444

 139 rProProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeuAspA 156
            ::::::|||   |||||||||   :::|||||||||:::|||:
 445 .AAAGGTGTGATACACAAGGACCTGAAGCCTGAGAATATCCTCGTTGATC 493

 156 snGluPheHisValLysIleAlaAspPheGlyLeuSerLysTrpArgMet 172
     :::::||||||:::||||||||||||:::|||:::::::::::::::
 494 GTGACTTTCACATTAAGATAGCCGATCTTGGTGTGGCTTCCTTTAAGACA 543

 173 MetSer...LeuSerGlnSerArgSerSerLysSerAlaPro........ 185
        |||   |||::::::::::::::::::|||:::
 544 TGGAGCAAACTGACTAAGGAGAAAGACAACAAGCAGAAAGAAGTGAGCAG 593

 186 ................GluGlyGlyThrIleIleTyrMetProProGlu. 196
                     :::|||||||||:::   ||||||   ||||||
 594 CACCACTAAGAAGAACAATGGTGGTACCCTTTACTACATGGCACCCGAAC 643

 197 ..............AsnTyrGluProGlyGlnLysSerArgAlaSerIle 208
                   |||   :::|||   :::||||||
 644 ACCTGAATGACATCAATGCAAAGCCCACGGAGAAGTCG............ 681

 209 LysHisAspIleTyrSerTyrAlaValIleThrTrpGluValLeuSerAr 225
           |||:::||||||::::::::::::   |||   :::::::::::
 682 ......GACGTGTACAGCTTTGGCATTGTCCTTTGGGCAATATTTGCAAA 725

 225 gLysGlnProPheGluAspValThrAsnProLeuGlnIleMetTyrSerV 242
     :|||:::|||:::|||:::|||            |||::::::      :
 726 AAAGGAGCCCTATGAGAATGTCATCTGTACTGAGCAGTTCGTGATCTGCA 775

 242 alSerGlnGlyHisArgProValIleAsnGlu......GluSerLeuPro 256
     ::::::::|||:::||||||   ::::::|||      |||      |||
 776 TAAAATCTGGGAACAGGCCAAATGTAGAGGAAATCCTTGAGTACTGTCCA 825

 257 TyrAspIleProHisArgAlaArgMetIleSerLeuIleGluSerGlyTr 273
        :::|||                  |||||||||:::|||      ||
 826 AGGGAGATC..................ATCAGCCTCATGGAGCGGTGCTG 857

 273 pAlaGlnAsnProAspGluArgProSerPheLeuLysCysLeuIleGluL 290
     |         |||::::::||||||:::||||||            |||:
 858 GCAGGCGATCCCAGAAGACAGGCCAACATTTCTTGGCATTGAAGAAGAAT 907

 290 euGluPro...ValLeuArgThrPheGluGluIleThrPheLeuGluAla 305
     :::::|||      |||      |||||||||   :::      |||
 908 TTAGGCCTTTTTACTTAAGTCATTTTGAAGAATATGTA...GAAGAGGAT 954

 306 ValIleGlnLeuLysLys..................ThrLysLeuGlnSe 316
     |||   :::|||||||||                        ||||||
 955 GTGGCAAGTTTAAAGAAAGAGTATCCAGATCAAAGCCCAGTGCTGCAGAG 1004

 316 rValSerSerAlaIleHisLeuCys......................... 324
      :::   |||      |||   |||
1005 AATGTTTTCACTGCAGCATGACTGTGTACCCTTACCTCCGAGCAGGTCAA 1054

 325 ..AspLysLysLysMetGluLeuSerLeuAsnIleProValAsnHisGly 340
       :::::::::      :::|||      :::      ::::::   |||
1055 ATTCAGAACAACCTGGATCGCTGCACAGTTCCCAGGGGCTCCAGATGGGT 1104

 341 ProGlnGluGluSerCysGlySerSerGln............LeuHisGl 353
     |||   |||||||||      ||||||:::                  ||
1105 CCTGTGGAGGAGTCCTGGTTTTCTTCCTCCCCAGAGTACCCACAGGACGA 1154

 353 uAsnSerGlySerProGluThr............................ 360
     ||||:::   |||   ::::::
1155 GAATGATCGCAGTGTGCAGGCTAAGCTGCAAGAGGAAGCCAGCTATCATG 1204

 361 ....................SerArgSerLeuProAlaProGlnAspAsn 370
                         :::::::::   |||   |||::::::|||
1205 CTTTTGGAATATTTGCAGAGAAACAGACAAAACCGCAGCCAAGGCAGAAT 1254

 371 AspPheLeuSerArgLysAlaGlnAspCysTyrPheMetLysLeuHisHi 387
     :::      :::|||:::   :::               ::::::   ||
1255 GAGGCTTACAACAGAGAGGAGGAA......AGGAAACGAAGGGTCTCTCA 1298

 387 sCysPro........................................... 389
     |   |||
1299 TGACCCCTTTGCACAGCAGAGAGCTCGTGAGAATATTAAGAGTGCAGGAG 1348

 390 ..GlyAsnHisSerTrpAspSerThrIleSerGlySerGlnArgAlaAla 405
          :::||||||      ||||||   |||         |||:::
1349 CAAGAGGTCATTCTGATCCCAGCACAACGAGT.........CGTGGAATT 1389

 406 PheCysAspHisLysThrThrProCysSerSerAlaIleIleAsn..... 420
           ::::::   :::   |||:::::::::::::::   |||
1390 GCAGTGCAACAGCTGTCATGGCCAGCCACCCAAACAGTTTGGAACAATGG 1439

 421 ................ProLeuSerThrAlaGlyAsnSerGluArgLeuG 432
                        ::::::|||:::|||::::::
1440 ATTGTATAATCAGCATGGATTTGGAACTACAGGTACAGGAGTTTGGTATC 1489

 432 lnProGlyIleAlaGlnGlnTrpIleGlnSerLys............... 443
       |||:::::::::|||::::::         |||
1490 CGCCAAATCTAAGCCAAATGTATAGTACTTATAAAACTCCAGTGCCAGAG 1539

 443 .................................................. 443

1540 ACCAACATACCGGGAAGCACACCCACCATGCCATACTTCTCTGGGCCAGT 1589

 444 .ArgGluAspIleVal.AsnGlnMet........................ 451
         :::|||:::::: |||   :::
1590 AGCAGATGACCTCATAAAATATACTATATTCAATAGTTCTGGTATTCAGA 1639

 452 ThrGluAlaCysLeuAsnGlnSerLeuAspAlaLeuLeuSerArgAspLe 468
        |||:::   :::         ||||||   :::   :::::::::
1640 TTGGAAACCACAATTATATGGATGTTGGACTGAATTCACAACCACCAAAC 1689

 468 uIleMetLysGluAspTyrGluLeuValSerThrLysProThrArgThrS 485
      |||:::   :::      :::|||   :::|||||||||:::      :
1690 AATACTTGCAAAGAAGAGTCGACTTCCAGACACCAAGCCATCTTTGATAA 1739

 485 erLysValArgGlnLeuLeuAspThrThrAspIleGlnGlyGlu...... 499
     ::   :::      |||::::::|||      ::::::|||:::
1740 CACCACTAGTCTGACTGATGAACACCTGAACCCTATCAGGGAAAACCTGG 1789

 500 ...........................GluPheAlaLysValIleValGl 507
                                ::::::|||::::::   :::::
1790 GAAGGCAGTGGAAAAACTGTGCCCGCAAGCTGGGCTTCACTGAGTCTCAG 1839

 507 nLysLeuLysAspAsnLysGlnMet 515
     :::::::|||::::::      |||
1840 ATCGATGAAATCGACCATGACTATG 1864
```

```
OM of: US-09-445-223-1 to: EST:*   out_format : pfs

Date: Jun 13, 2001  5:56 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn1_1/USPTO_spool/US09445223/runat_13062001_092032_29632/app_query.fasta_1.604
 -DB=EST -QFMT=fastap -SUFFIX=p2n.rst -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=2000000000 -USER=US09445223_@CGN1_1_3541 -NCPU=6
 -ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-445-223-1
Query length: 540
Database: EST:*
Database sequences: 9623517
Database length: 73081774
Search time (sec): 1225.860000


score_list:
Sequence            Strd Orig    ZScore   EScore Len  ! Documentation  ..
gb_est97:BG170405     + 1282.50 2325.49  1.9e-120  811  ! BG170405 602322736F1 NIH_MGC_89
gb_est79:BE877822     + 1168.00 2114.76  1.1e-108  870  ! BE877822 601486392F1 NIH_MGC_69
gb_est79:BE875947     + 1026.00 1855.18  3.0e-94   828  ! BE875947 601486423F1 NIH_MGC_69
gb_est75:BE536247     + 1019.00 1845.39  1.1e-93   654  ! BE536247 601062632F1 NIH_MGC_10
gb_est3:AA160647      +   995.00 1801.59  2.9e-91   645  ! AA160647 zq49c11.r1 Stratagene
gb_est24:AI745575     -   940.00 1701.91  1.0e-85   592  ! AI745575 wc34f12.x1 NCI_CGAP_Pr
gb_est53:AW960501     -   917.00 1658.62  2.7e-83   647  ! AW960501 EST372572 MAGE reseque
gb_est75:BE551615     -   894.00 1616.69  5.9e-81   636  ! BE551615 7a42g06.x1 NCI_CGAP_GC
gb_est3:AA161113      -   877.00 1587.00  2.6e-79   568  ! AA161113 zo58d05.s1 Stratagene
gb_est26:AI904799     +   876.00 1581.38  5.4e-79   762  ! AI904799 IL-BT067-190199-023 BT
gb_est72:BE274455     +   852.50 1535.93  1.8e-76   917  ! BE274455 601120495F1 NIH_MGC_20
gb_est41:AW085560     -   838.00 1514.69  2.8e-75   606  ! AW085560 wy67c04.x1 Soares_NSF_
gb_est5:AA315575      +   820.00 1484.61  1.3e-73   484  ! AA315575 EST187344 Colon carcin
gb_est10:AA655189     +   811.00 1467.62  1.2e-72   503  ! AA655189 vv13a12.r1 Stratagene
gb_est45:AW389863     -   798.50 1444.53  2.3e-71   510  ! AW389863 RC2-ST0173-041099-011-
gb_est92:BF796292     +   798.50 1436.65  6.3e-71   940  ! BF796292 602258601F1 NIH_MGC_85
gb_est91:BF706215     +   795.00 1439.12  4.6e-71   472  ! BF706215 280591 MARC 3BOV Bos t
gb_est45:AW361643     -   727.50 1315.70  3.4e-64   462  ! AW361643 MR2-CT0257-011199-011-
gb_est19:AI343247     -   713.00 1288.55  1.1e-62   483  ! AI343247 tb94a04.x1 NCI_CGAP_Co
gb_est13:AA913804     -   703.00 1270.42  1.1e-61   476  ! AA913804 ol35d11.s2 Soares_NFL_
gb_est9:AA574167      -   692.00 1251.78  1.2e-60   423  ! AA574167 nl78b01.s1 NCI_CGAP_Br
gb_est42:AW150819     -   684.00 1231.76  1.6e-59   641  ! AW150819 xg39f08.x1 NCI_CGAP_Ut
gb_est19:AI343850     -   669.00 1208.69  3.1e-58   455  ! AI343850 qp07d01.x1 NCI_CGAP_Ki
gb_est83:BF125028     +   665.00 1193.48  2.2e-57   839  ! BF125028 601762610F1 NIH_MGC_20
gb_est48:AW566614     +   663.00 1194.42  1.9e-57   587  ! AW566614 fk02e11.y2 zebrafish f
gb_est83:BF125423     +   648.00 1158.69  1.9e-55   1112  ! BF125423 601763549F1 NIH_MGC_2
gb_est89:BF565410     +   634.00 1144.22  1.2e-54   467  ! BF565410 UI-R-BO1-ajk-h-07-0-UI
gb_est25:AI801150     -   632.00 1137.01  3.1e-54   615  ! AI801150 to85h04.x1 NCI_CGAP_Ga
gb_est87:BF460494     +   630.00 1135.76  3.6e-54   510  ! BF460494 UI-M-CG0p-bmd-c-04-0-U
gb_est3:AA175489      +   629.00 1133.82  4.6e-54   514  ! AA175489 ms92a08.r1 Soares mous
gb_est48:AW593657     -   629.00 1133.72  4.7e-54   518  ! AW593657 xl94d07.x1 NCI_CGAP_Ut
gb_est16:AI123747     -   628.00 1133.91  4.6e-54   443  ! AI123747 oo21c04.x1 Soares_NSF_
gb_est45:AW389881     +   625.00 1130.66  6.9e-54   372  ! AW389881 RC4-ST0173-131099-031-
gb_est45:AW361603     +   624.00 1126.46  1.2e-53   447  ! AW361603 MR2-CT0257-011199-022-
gb_est46:AW446607     +   618.00 1116.64  4.2e-53   408  ! AW446607 86482 MARC 1BOV Bos ta
gb_est22:AI592519     +   614.00 1106.09  1.6e-52   524  ! AI592519 ms92a08.y1 Soares mous
gb_est11:AA723533     -   608.00 1097.58  4.8e-52   432  ! AA723533 zg73g08.s1 Soares_feta
gb_est46:AW465726     -   608.00 1097.00  5.2e-52   452  ! AW465726 BP230020A20D3 Soares n
gb_est18:AI307810     -   593.00 1065.52  2.9e-50   616  ! AI307810 tb28d07.x1 NCI_CGAP_Ki
gb_est105:W28134      -   591.00 1061.52  4.9e-50   632  ! W28134 42g10 Human retina cDNA
gb_est91:BF754627     +   579.00 1047.15  3.1e-49   350  ! BF754627 QV3-CT0558-121000-377-
gb_est5:AA301765      +   572.00 1034.81  1.5e-48   337  ! AA301765 EST14819 Aorta endo
gb_est68:BE002713     -   572.00 1031.20  2.4e-48   446  ! BE002713 QV4-BN0090-210300-1
gb_est12:AA826052     -   568.00 1025.27  5.1e-48   400  ! AA826052 od60b09.s1 NCI_CGAP
gb_est92:BF825562     -   565.00 1021.64  8.2e-48   346  ! BF825562 CM1-HN0016-161100-5

seq_name: gb_est97:BG170405

seq_documentation_block:
LOCUS       BG170405      811 bp    mRNA            EST       06-FEB-2001
DEFINITION  602322736F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4426016 5',
            mRNA sequence.
ACCESSION   BG170405
VERSION     BG170405.1  GI:12677108
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 811)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10172  row: g  column: 09
            High quality sequence stop: 721.
FEATURES             Location/Qualifiers
     source          1..811
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4426016"
                     /clone_lib="NIH_MGC_89"
                     /tissue_type="hypernephroma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      272 a    177 c    156 g    206 t
ORIGIN


alignment_scores:
          Quality: 1282.50              Length:    270
            Ratio:  4.933                 Gaps:      3
 Percent Similarity:  96.296   Percent Identity: 94.074

alignment_block:
 US-09-445-223-1 x BG170405   ..

 Align seg 1/1  to: BG170405  from: 1  to: 811

     259 IleProHisArgAlaArgMetIleSerLeuIleGluSerGlyTrpAlaGl 275
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       3 ATACCTCACCGAGCACGTATGATCTCTCTAATAGAAAGTGGATGGGCACA 52

     275 nAsnProAspGluArgProSerPheLeuLysCysLeuIleGluLeuGluP 292
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      53 AAATCCAGATGAAAGACCATCTTTCTTAAAATGTTTAATAGAACTTGAAC 102

     292 roValLeuArgThrPheGluGluIleThrPheLeuGluAlaValIleGln 308
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     103 CAGTTTTGAGAACATTTGAAGAGATAACTTTTCTTGAAGCTGTTATTCAG 152
```

```
309 LeuLysLysThrLysLeuGlnSerValSerSerAlaIleHisLeuCysAs 325
    ||||||||||||||||||||||||||||||||||||||||||||||||||
153 CTAAAGAAAACAAAGTTACAGAGTGTTTCAAGTGCCATTCACCTATGTGA 202

325 pLysLysLysMetGluLeuSerLeuAsnIleProValAsnHisGlyProG 342
    ||||||||||||||||||||||||||||||||||||||||||||||||||
203 CAAGAAGAAAATGGAATTATCTCTGAACATACCTGTAAATCATGGTCCAC 252

342 lnGluGluSerCysGlySerSerGlnLeuHisGluAsnSerGlySerPro 358
    ||||||||||||||||||||||||||||||||||||||||||||||||||
253 AAGAGGAATCATGTGGATCCTCTCAGCTCCATGAAAATAGTGGTTCTCCT 302

359 GluThrSerArgSerLeuProAlaProGlnAspAsnAspPheLeuSerAr 375
    ||||||||||||||||||||||||||||||||||||||||||||||||||
303 GAAACTTCAAGGTCCCTGCCAGCTCCTCAAGACAATGATTTTTTATCTAG 352

375 gLysAlaGlnAspCysTyrPheMetLysLeuHisHisCysProGlyAsnH 392
    ||||||||||||||||||||||||||||||||||||||||||||||||||
353 AAAAGCTCAAGACTGTTATTTTATGAAGCTGCATCACTGTCCTGGAAATC 402

392 isSerTrpAspSerThrIleSerGlySerGlnArgAlaAlaPheCysAsp 408
    ||||||||||||||||||||||||||||||||||||||||||||||||||
403 ACAGTTGGGATAGCACCATTTCTGGATCTCAAAGGGCTGCATTCTGTGAT 452

409 HisLysThrThrProCysSerSerAlaIleIleAsnProLeuSerThrAl 425
    ||||||||||||||||||||||||||||||||||||||||||||||||||
453 CACAAGACCACTCCATGCTCTTCAGCAATAATAAATCCACTCTCAACTGC 502

425 aGlyAsnSerGluArgLeuGlnProGlyIleAlaGlnGlnTrpIleGlnS 442
    ||||||||||||||||||||||||||||||||||||||||||||||||||
503 AGGAAACTCAGAACGTCTGCAGCCTGGTATAGCCCAGCAGTGGATCCAGA 552

442 erLysArgGluAspIleValAsnGlnMetThrGluAlaCysLeuAsnGln 458
    ||||||||||||||||||||||||||||||||||||||||||||||||||
553 GCAAAAGGGAAGACATTGTGAACCAAATGACAGAAGCCTGCCTTAACCAG 602

459 SerLeuAspAlaLeuLeuSerArgAspLeuIleMetLysGluAspTyrGl 475
    ||||||||||||||||||||||||||||||||||||||||||||||||||
603 TCGCTAGATGCCCTTCTGTCCAGGGACTTGATCATGAAAGAGGACTATGA 652

475 uLeuValSerThrLysProThrArgThrSerLysValArg.GlnLeuLeu 491
    |||||||||||||||||||||||||||||||||||||||| |||||||||
653 ACTTGTTAGTACCAAGCCTACAAGGACCTCAAAAGTCAGACCAATTACTA 702

492 AspThrThrAspIleGlnGlyGluGluPheAlaLysValIleValGlnLy 508
    ||||||||||||||||||||||||::::::   |||:::   ||||||||
703 GACACTACTGACATCCAAGGAGAAGATTTGCCAAAGTTATACGTACAAAA 752

508 sLeu.LysAspAsnLysGlnMetGlyLeuGlnProTyrProGluIleLeu 524
    |||| :::            |||||||||:::||||||  ':::||||||
753 ATTGACGA...TTACCAACAATGGGTCTTAAGCCTTACCGGAAAATACTG 799

525 ValVal 526
       |||
800 TGGGTT 805

seq_name: gb_est79:BE877822

seq_documentation_block:
LOCUS       BE877822      870 bp    mRNA            EST       20-OCT-2000
DEFINITION  601486392F1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:3888815 5',
            mRNA sequence.
ACCESSION   BE877822
VERSION     BE877822.1  GI:10326598
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 870)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
```

```
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9669  row: g  column: 24
            High quality sequence stop: 739.
FEATURES             Location/Qualifiers
     source          1..870
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3888815"
                     /clone_lib="NIH_MGC_69"
                     /tissue_type="large cell carcinoma, undifferentiated"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.1 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      288 a    200 c    170 g    212 t
ORIGIN


alignment_scores:
           Quality: 1168.00              Length:    244
             Ratio:  4.846                 Gaps:      3
 Percent Similarity:  98.770   Percent Identity: 98.361

alignment_block:
 US-09-445-223-1 x BE877822   ..

 Align seg 1/1  to: BE877822  from: 1  to: 870

     299 GluIleThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGl 315
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       2 GAGATAACTTTTCTTGAAGCTGTTATTCAGCTAAAGAAAACAAAGTTACA 51

     315 nSerValSerSerAlaIleHisLeuCysAspLysLysLysMetGluLeuS 332
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      52 GAGTGTTTCAAGTGCCATTCACCTATGTGACAAGAAGAAAATGGAATTAT 101

     332 erLeuAsnIleProValAsnHisGlyProGlnGluGluSerCysGlySer 348
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     102 CTCTGAACATACCTGTAAATCATGGTCCACAAGAGGAATCATGTGGATCC 151

     349 SerGlnLeuHisGluAsnSerGlySerProGluThrSerArgSerLeuPr 365
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     152 TCTCAGCTCCATGAAAATAGTGGTTCTCCTGAAACTTCAAGGTCCCTGCC 201

     365 oAlaProGlnAspAsnAspPheLeuSerArgLysAlaGlnAspCysTyrP 382
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     202 AGCTCCTCAAGACAATGATTTTTTATCTAGAAAAGCTCAAGACTGTTATT 251

     382 heMetLysLeuHisHisCysProGlyAsnHisSerTrpAspSerThrIle 398
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     252 TTATGAAGCTGCATCACTGTCCTGGAAATCACAGTTGGGATAGCACCATT 301

     399 SerGlySerGlnArgAlaAlaPheCysAspHisLysThrThrProCysSe 415
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     302 TCTGGATCTCAAAGGGCTGCATTCTGTGATCACAAGACCACTCCATGCTC 351

     415 rSerAlaIleIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeuG 432
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     352 TTCAGCAATAATAAATCCACTCTCAACTGCAGGAAACTCAGAACGTCTGC 401

     432 lnProGlyIleAlaGlnGlnTrpIleGlnSerLysArgGluAspIleVal 448
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     402 AGCCTGGTATAGCCCAGCAGTGGATCCAGAGCAAAAGGGAAGACATTGTG 451
```

```
    449 AsnGlnMetThrGluAlaCysLeuAsnGlnSerLeuAspAlaLeuLeuSe 465
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    452 AACCAAATGACAGAAGCCTGCCTTAACCAGTCGCTAGATGCCCTTCTGTC 501

    465 rArgAspLeuIleMetLysGluAspTyrGluLeuValSerThrLysProT 482
        |   ||||||||||||||||||||||||||||||||||||||||||||||
    502 CAG.GACTTGATCATGAAAGAGGACTATGAACTTGTTAGTACCAAGCCTA 550

    482 hrArgThrSerLysValArgGlnLeuLeuAspThrThrAspIleGln.Gl 498
        ||||||||||||||||||||||||||||||||||||||||||||||| ||
    551 CAAGGACCTCAAAAGTCAGACAATTACTAGACACTACTGACATCCAAGGG 600

    498 yGluGluPheAlaLysValIleValGlnLysLeuLysAspAsnLysGlnM 515
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    601 AGAAGAATTTGCCAAAGTTATAGTACAAAAATTGAAAGATAACAAACAAA 650

    515 etGlyLeu..GlnProTyrProGluIleLeuValValSerArgSerProS 531
        ||||||||   :::||||||||||||||||||||||||||||||||||||
    651 TGGGTCTTTCAGGCCTTACCCGGAAATACTTGTGGTTTCTAGATCACCAT 700

    531 erLeuAsn.LeuLeuGlnAsnLysSer 539
        |||||||| ||||||||||||||||||
    701 CTTTAAATTTTACTTCAGAATAAAGC 727

seq_name: gb_est79:BE875947

seq_documentation_block:
LOCUS       BE875947       828 bp    mRNA            EST       20-OCT-2000
DEFINITION  601486423F1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:3888965 5',
            mRNA sequence.
ACCESSION   BE875947
VERSION     BE875947.1  GI:10324723
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 828)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9669  row: n  column: 06
            High quality sequence stop: 795.
FEATURES             Location/Qualifiers
     source          1..828
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3888965"
                     /clone_lib="NIH_MGC_69"
                     /tissue_type="large cell carcinoma, undifferentiated"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.1 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      287 a    163 c    165 g    213 t
ORIGIN


alignment_scores:
            Quality: 1026.00             Length:    205
              Ratio:  5.029                Gaps:      1
```

```
 Percent Similarity:  99.512   Percent Identity: 99.512

alignment_block:
 US-09-445-223-1 x BE875947   ..

 Align seg 1/1  to: BE875947  from: 1  to: 828

    337 ValAsnHisGlyProGlnGluGluSerCysGlySerSerGlnLeuHisGl 353
        ||||||||||||||||||||||||||||||||||||||||||||||||||
      5 GTAAATCATGGTCCACAAGAGGAATCATGTGGATCCTCTCAGCTCCATGA 54

    353 uAsnSerGlySerProGluThrSerArgSerLeuProAlaProGlnAspA 370
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     55 AAATAGTGGTTCTCCTGAAACTTCAAGGTCCCTGCCAGCTCCTCAAGACA 104

    370 snAspPheLeuSerArgLysAlaGlnAspCysTyrPheMetLysLeuHis 386
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    105 ATGATTTTTATCTAGAAAAGCTCAAGACTGTTATTTATGAAGCTGCAT 154

    387 HisCysProGlyAsnHisSerTrpAspSerThrIleSerGlySerGlnAr 403
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    155 CACTGTCCTGGAAATCACAGTTGGGATAGCACCATTTCTGGATCTCAAAG 204

    403 gAlaAlaPheCysAspHisLysThrThrProCysSerSerAlaIleIleA 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    205 GGCTGCATTCTGTGATCACAAGACCACTCCATGCTCTTCAGCAATAATAA 254

    420 snProLeuSerThrAlaGlyAsnSerGluArgLeuGlnProGlyIleAla 436
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    255 ATCCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAGCCTGGTATAGCC 304

    437 GlnGlnTrpIleGlnSerLysArgGluAspIleValAsnGlnMetThrGl 453
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    305 CAGCAGTGGATCCAGAGCAAAAGGGAAGACATTGTGAACCAAATGACAGA 354

    453 uAlaCysLeuAsnGlnSerLeuAspAlaLeuLeuSerArgAspLeuIleM 470
        |||||||||||||||||||||||||||||||||||||   ||||||||||
    355 AGCCTGCCTTAACCAGTCGCTAGATGCCCTTCTGTCCAG.GACTTGATCA 403

    470 etLysGluAspTyrGluLeuValSerThrLysProThrArgThrSerLys 486
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    404 TGAAAGAGGACTATGAACTTGTTAGTACCAAGCCTACAAGGACCTCAAAA 453

    487 ValArgGlnLeuLeuAspThrThrAspIleGlnGlyGluGluPheAlaLy 503
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    454 GTCAGACAATTACTAGACACTACTGACATCCAAGGAGAAGAATTTGCCAA 503

    503 sValIleValGlnLysLeuLysAspAsnLysGlnMetGlyLeuGlnProT 520
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    504 AGTTATAGTACAAAAATTGAAAGATAACAAACAAATGGGTCTTCAGCCTT 553

    520 yrProGluIle.LeuValValSerArgSerProSerLeuAsnLeuLeuGl 536
        |||||||||||| |||||||||||||||||||||||||||||||||||||
    554 ACCCGGAAATACTTGGTGGTTTCTAGATCACCATCTTTAAATTTACTTCA 603

    536 nAsnLysSerMet 540
        |||||||||||||
    604 AAATAAAAGCATG 616

seq_name: gb_est75:BE536247

seq_documentation_block:
LOCUS       BE536247       654 bp    mRNA            EST       09-AUG-2000
DEFINITION  601062632F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:3449311 5',
            mRNA sequence.
ACCESSION   BE536247
VERSION     BE536247.1  GI:9764892
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 654)
```

```
  AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
  TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL    Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Tel: (301) 496-1550
             Email: Robert_Strausberg@nih.gov
             Tissue Procurement: ATCC
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: Incyte Genomics, Inc.
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLAM8425  row: o  column: 08
             High quality sequence stop: 650.
FEATURES             Location/Qualifiers
     source          1. .654
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3449311"
                     /clone_lib="NIH_MGC_10"
                     /cell_line="MGC36"
                     /lab_host="DH10B"
                     /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.5 kb. Library prepared by Life
                     Technologies."
BASE COUNT      214 a    149 c    131 g    158 t      2 others
ORIGIN


alignment_scores:
           Quality: 1019.00           Length:    219
             Ratio:  4.807              Gaps:      4
 Percent Similarity:  96.804   Percent Identity: 95.890

alignment_block:
 US-09-445-223-1 x BE536247   ..

 Align seg 1/1  to: BE536247  from: 1  to: 654

     311 LysThrLysLeuGlnSerValSerSerAlaIleHisLeuCysAspLysLy 327
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       3 AAAACAAAGTTACAGAGTGTTTCAAGTGCCATTCACCTATGTGACAAGAA 52

     327 sLysMetGluLeuSerLeuAsnIleProValAsnHisGlyProGlnGluG 344
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      53 GAAAATGGAATTATCTCTGAACATACCTGTAAATCATGGTCCACAAGAGG 102

     344 luSerCysGlySerSer.GlnLeuHisGluAsnSerGlySerProGluTh 360
         |||||||||||||||||    |||||||||||||||||||||||||||||
     103 AATCATGTGGATCCTCTCANGCTCCATGAAAATAGTGGTTCTCCTGAAAC 152

     360 rSerArgSerLeuProAlaProGlnAspAsnAspPheLeuSerArgLysA 377
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     153 TTCAAGGTCCCTGCCAGCTCCTCAAGACAATGATTTTTATCTAGAAAAG 202

     377 laGlnAspCysTyrPheMetLysLeuHisHisCysProGlyAsnHisSer 393
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     203 CTCAAGACTGTTATTTTATGAAGCTGCATCACTGTCCTGGAAATCACAGT 252

     394 TrpAspSerThrIleSerGlySerGlnArgAlaAlaPheCysAspHisLy 410
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     253 TGGGATAGCACCATTTCTGGATCTCAAAGGGCTGCATTCTGTGATCACAA 302

     410 sThrThrProCysSerSerAlaIleIleAsnProLeuSerThrAlaGlyA 427
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     303 GACCACTCCATGCTCTTCAGCAATAATAAATCCACTCTCAACTGCAGGAA 352

     427 snSerGluArgLeuGlnProGlyIleAlaGlnGlnTrpIleGlnSerLys 443
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     353 ACTCAGAACGTCTGCAGCCTGGTATAGCCCAGCAGTGGATCCAGAGCAAA 402
```

```
     444 ArgGluAspIleValAsnGlnMetThrGluAlaCysLeuAsnGlnSerLe 460
         ||||||||||||||||||||||||||||||||||||   |||||||||||
     403 AGGGAAGACATTGTGAACCAAATGACAGAAGCCTGCTT.AACCAGTCGCT 451

     460 uAspAlaLeuLeuSerArgAspLeuIleMetLysGluAspTyrGluLeuV 477
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     452 AGATGCCCTTCTGTCCAGGGACTTGATCATGAAAGAGGACTATGAACTTG 501

     477 alSerThrLysProThrArgThrSerLysValArgGlnLeuLeuAspThr 493
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     502 TTAGTACCAAGCCTACAAGGACCTCAAAAGTCAGACAATTGCTAGACACT 551

     494 ThrAspIleGlnGly.GluGluPheAlaLysValIleValGlnLys.Leu 509
         ||||||||||||||| |||||||||:::   ||||||||||||||| |||
     552 ACTGACATCCAAGGCAGAAGAATTTGGCANAGTTATAGTACAAAAATTTG 601

     510 LysAspAsn.LysGlnMetGlyLeuGlnProTyrProGluIleLeuVal 525
         :::|||||| |||||||||||||||||||||||||||   |||||||||
     602 CAAGATAACCAAACAAATGGGTCTTCAGCCTTACCCGGGAATACTTGTG 650

seq_name: gb_est3:AA160647

seq_documentation_block:
LOCUS       AA160647      645 bp    mRNA            EST       16-DEC-1996
DEFINITION  zq49c11.r1 Stratagene hNT neuron (#937233) Homo sapiens cDNA clone
            IMAGE:633044 5', mRNA sequence.
ACCESSION   AA160647
VERSION     AA160647.1  GI:1736024
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 645)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 373.
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="Homo sapiens"
                     /db_xref="GDB:5183128"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:633044"
                     /clone_lib="Stratagene hNT neuron (#937233)"
                     /dev_stage="hNT neurons"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT.
                     Differentiated, post mitotic hNT neurons. Average insert
                     size: 1.5 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5'
                     GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTTTT 3'"
BASE COUNT      205 a    138 c    123 g    178 t      1 others
ORIGIN


alignment_scores:
```

```
          Quality:  995.00           Length:    211
            Ratio:  4.901              Gaps:      2
 Percent Similarity:  96.209   Percent Identity: 92.891

alignment_block:
 US-09-445-223-1 x AA160647   ..

 Align seg 1/1  to: AA160647  from: 1  to: 645

     232 ValThrAsnProLeuGlnIleMetTyrSerValSerGlnGlyHisArgPr 248
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       2 GTCACCAATCCTTTGCAGATAATGTATAGTGTGTCACAAGGACATCGACC 51

     248 oValIleAsnGluGluSerLeuProTyrAspIleProHisArgAlaArgM 265
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      52 TGTTATTAATGAAGAAAGTTTGCCATATGATATACCTCACCGAGCACGTA 101

     265 etIleSerLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgPro 281
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     102 TGATCTCTCTAATAGAAAGTGGATGGGCACAAAATCCAGATGAAAGACCA 151

     282 SerPheLeuLysCysLeuIleGluLeuGluProValLeuArgThrPheGl 298
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     152 TCTTTCTTAAAATGTTTAATAGAACTTGAACCAGTTTTGAGAACATTTGA 201

     298 uGluIleThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuG 315
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     202 AGAGATAACTTTTCTTGAAGCTGTTATTCAGCTAAAGAAAACAAAGTTAC 251

     315 lnSerValSerSerAlaIleHisLeuCysAspLysLysLysMetGluLeu 331
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     252 AGAGTGTTTCAAGTGCCATTCACCTATGTGACAAGAAGAAAATGGAATTA 301

     332 SerLeuAsnIleProValAsnHisGlyProGlnGluGluSerCysGlySe 348
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     302 TCTCTGAACATACCTGTAAATCATGGTCCACAAGAGGAATCATGTGGATC 351

     348 rSerGlnLeuHisGluAsnSerGlySerProGluThrSerArgSerLeuP 365
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     352 CTCTCAGCTCCATGAAAATAGTGGTTCTCCTGAAACTTCAAGGTCCCTGC 401

     365 roAlaProGlnAspAsnAspPheLeuSerArgLysAlaGlnAspCysTyr 381
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     402 CAGCTCCTCAAGACAATGATTTTTATCTAGAAAAGCTCAAGACTGTTAT 451

     382 PheMetLysLeuHisHisCysProGlyAsnHisSerTrpAspSerThrIl 398
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     452 TTTATGAAGCTGCATCACTGTCCTGGAAATCACAGTTGGGATAGCACCAT 501

     398 eSerGlySerGlnArgAlaAlaPheCysAspHisLysThrThrProCysS 415
         |||||||||||||||||||||||||||||||   :::|||||||||||||
     502 TTCTGGATCTCAAAGGGCTGCATTCTGTGATTCACAGACCACTCCATGCT 551

     415 erSerAlaIleIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeu 431
         |||||   ||||||||||||||||||         :::   :::   |||
     552 CTTCAGC.ATAATAAATCCACTCTCACTGCAG...GAACTCCGAACGTTG 597

     432 GlnProGlyIleAlaGlnGlnTrpIleGlnSer 442
         ||||||   ::::::   |||||||||::::::
     598 CAGCCT...GTATGCCCGCAGTGGATCCGAGCA 627

seq_name: gb_est24:AI745575

seq_documentation_block:
LOCUS       AI745575      592 bp    mRNA            EST       17-DEC-1999
DEFINITION  wc34f12.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2317103 3'
            similar to TR:O43353 O43353 SERINE/THREONINE KINASE RICK. ;, mRNA
            sequence.
ACCESSION   AI745575
VERSION     AI745575.1  GI:5113863
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 592)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 673   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 459.
FEATURES             Location/Qualifiers
     source          1. .592
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2317103"
                     /clone_lib="NCI_CGAP_Pr28"
                     /sex="male"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Plasmid DNA from the
                     normalized library NCI_CGAP_Pr22 was prepared, and ss
                     circles were made in vitro. Following HAP purification,
                     this DNA was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from a pool
                     of 5,000 clones made from the same library (cloneIDs
                     985608-986759, 1101192-1101959, and 1217928-1220615).
                     Subtraction by Bento Soares and M. Fatima Bonaldo.  "
BASE COUNT      153 a    111 c    128 g    200 t
ORIGIN


alignment_scores:
            Quality:  940.00             Length:    181
              Ratio:  5.193                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-445-223-1 x AI745575/rev   ..

 Align seg 1/1  to reverse of: AI745575  from: 1  to: 592

     360 ThrSerArgSerLeuProAlaProGlnAspAsnAspPheLeuSerArgLy 376
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     592 ACTTCAAGGTCCCTGCCAGCTCCTCAAGACAATGATTTTTATCTAGAAA 543

     376 sAlaGlnAspCysTyrPheMetLysLeuHisHisCysProGlyAsnHisS 393
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     542 AGCTCAAGACTGTTATTTTATGAAGCTGCATCACTGTCCTGGAAATCACA 493

     393 erTrpAspSerThrIleSerGlySerGlnArgAlaAlaPheCysAspHis 409
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     492 GTTGGGATAGCACCATTTCTGGATCTCAAAGGGCTGCATTCTGTGATCAC 443

     410 LysThrThrProCysSerSerAlaIleIleAsnProLeuSerThrAlaGl 426
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     442 AAGACCACTCCATGCTCTTCAGCAATAATAAATCCACTCTCAACTGCAGG 393

     426 yAsnSerGluArgLeuGlnProGlyIleAlaGlnGlnTrpIleGlnSerL 443
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     392 AAACTCAGAACGTCTGCAGCCTGGTATAGCCCAGCAGTGGATCCAGAGCA 343
```

```
    443 ysArgGluAspIleValAsnGlnMetThrGluAlaCysLeuAsnGlnSer 459
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    342 AAAGGGAAGACATTGTGAACCAAATGACAGAAGCCTGCCTTAACCAGTCG 293

    460 LeuAspAlaLeuLeuSerArgAspLeuIleMetLysGluAspTyrGluLe 476
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    292 CTAGATGCCCTTCTGTCCAGGGACTTGATCATGAAAGAGGACTATGAACT 243

    476 uValSerThrLysProThrArgThrSerLysValArgGlnLeuLeuAspT 493
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    242 TGTTAGTACCAAGCCTACAAGGACCTCAAAAGTCAGACAATTACTAGACA 193

    493 hrThrAspIleGlnGlyGluGluPheAlaLysValIleValGlnLysLeu 509
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    192 CTACTGACATCCAAGGAGAAGAATTTGCCAAAGTTATAGTACAAAAATTG 143

    510 LysAspAsnLysGlnMetGlyLeuGlnProTyrProGluIleLeuValVa 526
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    142 AAAGATAACAAACAAATGGGTCTTCAGCCTTACCCGGAAATACTTGTGGT 93

    526 lSerArgSerProSerLeuAsnLeuLeuGlnAsnLysSerMet 540
        |||||||||||||||||||||||||||||||||||||||||||
     92 TTCTAGATCACCATCTTTAAATTTACTTCAAAATAAAAGCATG 50

seq_name: gb_est53:AW960501

seq_documentation_block:
LOCUS       AW960501      647 bp    mRNA             EST       01-JUN-2000
DEFINITION  EST372572 MAGE resequences, MAGF Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW960501
VERSION     AW960501.1  GI:8150185
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 647)
  AUTHORS   Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt
            ,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
            Quackenbush,J.
  TITLE     Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
  JOURNAL   Unpublished (2000)
COMMENT     Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            Plate: 147
            Seq primer: Forward.
FEATURES             Location/Qualifiers
     source          1..647
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="MAGE resequences, MAGF"
                     /note="Vector: pBluescriptSKm"
BASE COUNT      165 a    128 c    141 g    213 t
ORIGIN


alignment_scores:
            Quality:  917.00              Length:    196
              Ratio:  4.930                 Gaps:      1
 Percent Similarity:  94.898   Percent Identity: 92.857

alignment_block:
 US-09-445-223-1 x AW960501/rev   ..

 Align seg 1/1  to reverse of: AW960501  from: 1  to: 647

     346 CysGlySerSerGlnLeuHisGluAsnSerGlySerProGluThrSerAr 362
         |||   :::         ::::::||||||||||||   :::   |||||
```

```
    634 TGCCAGAATTTCGGCACGAGGAAAAATAGTGGTTCTCTTAAAATTTCAAG 585

    362 g.SerLeuProAlaProGlnAspAsnAspPheLeuSerArgLysAlaGln 378
        | ||||||||||||   |||   |||||||||||||||||||||||||||
    584 GGTCCCTGCCAGCTCTCCAAGCCAATGATTTTTATCTAGAAAAGCTCAA 535

    379 AspCysTyrPheMetLysLeuHisHisCysProGlyAsnHisSerTrpAs 395
           |||||||||||||||||||||||||||||||||||||||||||||||
    534 GCCTGTTATTTATGAAGCTGCATCACTGTCCTGGAAATCACAGTTGGGA 485

    395 pSerThrIleSerGlySerGlnArgAlaAlaPheCysAspHisLysThrT 412
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    484 TAGCACCATTTCTGGATCTCAAAGGGCTGCATTCTGTGATCACAAGACCA 435

    412 hrProCysSerSerAlaIleIleAsnProLeuSerThrAlaGlyAsnSer 428
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    434 CTCCATGCTCTTCAGCAATAATAAATCCACTCTCAACTGCAGGAAACTCA 385

    429 GluArgLeuGlnProGlyIleAlaGlnGlnTrpIleGlnSerLysArgGl 445
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    384 GAACGTCTGCAGCCTGGTATAGCCCAGCAGTGGATCCAGAGCAAAAGGGA 335

    445 uAspIleValAsnGlnMetThrGluAlaCysLeuAsnGlnSerLeuAspA 462
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    334 AGACATTGTGAACCAAATGACAGAAGCCTGCCTTAACCAGTCGCTAGATG 285

    462 laLeuLeuSerArgAspLeuIleMetLysGluAspTyrGluLeuValSer 478
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    284 CCCTTCTGTCCAGGGACTTGATCATGAAAGAGGACTATGAACTTGTTAGT 235

    479 ThrLysProThrArgThrSerLysValArgGlnLeuLeuAspThrThrAs 495
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    234 ACCAAGCCTACAAGGACCTCAAAAGTCAGACAATTACTAGACACTACTGA 185

    495 pIleGlnGlyGluGluPheAlaLysValIleValGlnLysLeuLysAspA 512
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    184 CATCCAAGGAGAAGAATTTGCCAAAGTTATAGTACAAAAATTGAAAGATA 135

    512 snLysGlnMetGlyLeuGlnProTyrProGluIleLeuValValSerArg 528
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    134 ACAAACAAATGGGTCTTCAGCCTTACCCGGAAATACTTGTGGTTTCTAGA 85

    529 SerProSerLeuAsnLeuLeuGlnAsnLysSerMet 540
        ||||||||||||||||||||||||||||||||||||
     84 TCACCATCTTTAAATTTACTTCAAAATAAAAGCATG 49

seq_name: gb_est75:BE551615

seq_documentation_block:
LOCUS       BE551615      636 bp    mRNA             EST       10-AUG-2000
DEFINITION  7a42g06.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:3221434 3'
            similar to TR:O43353 O43353 SERINE/THREONINE KINASE RICK. ;, mRNA
            sequence.
ACCESSION   BE551615
VERSION     BE551615.1  GI:9793227
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 636)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
```

```
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 455.
FEATURES             Location/Qualifiers
     source          1..636
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3221434"
                     /clone_lib="NCI_CGAP_GC6"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Plasmid DNA from the normalized library
                     NCI_CGAP_GC4 was prepared, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from a pool of 5,000 clones made
                     from the same library (cloneIDs 1257096-1258631,
                     1469064-1470983, and 1475592-1476743). Subtraction by
                     Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      175 a    118 c    135 g    208 t
ORIGIN


alignment_scores:
           Quality:  894.00             Length:    181
             Ratio:  5.022                Gaps:      0
 Percent Similarity:  98.343   Percent Identity: 98.343

alignment_block:
 US-09-445-223-1 x BE551615/rev   ..

 Align seg 1/1  to reverse of: BE551615  from: 1  to: 636

     360 ThrSerArgSerLeuProAlaProGlnAspAsnAspPheLeuSerArgLy 376
         ||||||   ||||||||||||||||||||||||   ||||||||||||||
     636 ACTTCAGGGTCCCTGCCAGCTCCTCAAGACAATGT.TTTTTATCTAGAAA 588

     376 sAlaGlnAspCysTyrPheMetLysLeuHisHisCysProGlyAsnHisS 393
         |||||||||||||||||||||||||||||||   ||||||||||||||||
     587 AGCTCAAGACTGTTATTTTATGAAGCTGCATCCCTGTCCTGGAAATCACA 538

     393 erTrpAspSerThrIleSerGlySerGlnArgAlaAlaPheCysAspHis 409
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     537 GTTGGGATAGCACCATTTCTGGATCTCAAAGGGCTGCATTCTGTGATCAC 488

     410 LysThrThrProCysSerSerAlaIleIleAsnProLeuSerThrAlaGl 426
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     487 AAGACCACTCCATGCTCTTCAGCAATAATAAATCCACTCTCAACTGCAGG 438

     426 yAsnSerGluArgLeuGlnProGlyIleAlaGlnGlnTrpIleGlnSerL 443
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     437 AAACTCAGAACGTCTGCAGCCTGGTATAGCCCAGCAGTGGATCCAGAGCA 388

     443 ysArgGluAspIleValAsnGlnMetThrGluAlaCysLeuAsnGlnSer 459
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     387 AAAGGAAGACATTGTGAACCAAATGACAGAAGCCTGCCTTAACCAGTCG 338

     460 LeuAspAlaLeuLeuSerArgAspLeuIleMetLysGluAspTyrGluLe 476
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     337 CTAGATGCCCTTCTGTCCAGGGACTTGATCATGAAAGAGGACTATGAACT 288

     476 uValSerThrLysProThrArgThrSerLysValArgGlnLeuLeuAspT 493
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     287 TGTTAGTACCAAGCCTACAAGGACCTCAAAAGTCAGACAATTACTAGACA 238

     493 hrThrAspIleGlnGlyGluGluPheAlaLysValIleValGlnLysLeu 509
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     237 CTACTGACATCCAAGGAGAAGAATTTGCCAAAGTTATAGTACAAAAATTG 188
```

```
     510 LysAspAsnLysGlnMetGlyLeuGlnProTyrProGluIleLeuValVa 526
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     187 AAAGATAACAAACAAATGGGTCTTCAGCCTTACCCGGAAATACTTGTGGT 138

     526 lSerArgSerProSerLeuAsnLeuLeuGlnAsnLysSerMet 540
         |||||||||||||||||||||||||||||||||||||||||||
     137 TTCTAGATCACCATCTTTAAATTTACTTCAAAATAAAAGCATG 95

seq_name: gb_est3:AA161113

seq_documentation_block:
LOCUS       AA161113      568 bp    mRNA            EST       09-MAR-1998
DEFINITION  zo58d05.s1 Stratagene pancreas (#937208) Homo sapiens cDNA clone
            IMAGE:591081 3', mRNA sequence.
ACCESSION   AA161113
VERSION     AA161113.1  GI:1735349
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 568)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1942   Std Error: 0.00
            Seq primer: -40M13 fwd. from Amersham
            High quality sequence stop: 256.
FEATURES             Location/Qualifiers
     source          1..568
                     /organism="Homo sapiens"
                     /db_xref="GDB:4622467"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:591081"
                     /clone_lib="Stratagene pancreas (#937208)"
                     /lab_host="SOLR cells (kanamycin resistant)"
                     /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. Pancreatic adenocarcinoma cell line. Average
                     insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor
                     sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTTTT 3'"
BASE COUNT      140 a    111 c    127 g    185 t      5 others
ORIGIN


alignment_scores:
           Quality:  877.00             Length:    183
             Ratio:  4.983                Gaps:      2
 Percent Similarity:  96.175   Percent Identity: 95.628

alignment_block:
 US-09-445-223-1 x AA161113/rev   ..

 Align seg 1/1  to reverse of: AA161113  from: 1  to: 568

     360 ThrSerArgSerLeuProAlaProGlnAspAsnAspPheLeuSerArgLy 376
         ||||||   |||||||||||||||||||||||||||||||||||||||||
     568 ACTTCAGGGTCCCTGCCAGCTCCTCAAGACAATGATTTTTTATCTAGAAA 519

     376 sAlaGlnAspCysTyrPheMetLysLeuHisHisCysProGlyAsnHis. 392
         |||||||||||||||||||   |||||||||||||||||||||||||||
```

```
    518 GGCTCAAGACTGTTATTTANGAAGCTGCATCACTGTCCTGGAAATCACC 469

    393 SerTrpAspSerThrIleSerGlySerGlnArgAlaAlaPheCysAspHi 409
        ||||||||||||||||||||||||||||||   |||:::|||||||||||
    468 AGTTGGGATAGCACCATTTCTGGATCTCAANGGGCTNCATTCTGTGATCA 419

    409 sLysThrThrProCysSerSerAlaIleIleAsnProLeuSerThrAlaG 426
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    418 CAAGACCACTCCATGCTCTTCAGCAATAATAAATCCACTCTCAACTGCAG 369

    426 lyAsnSerGluArgLeuGlnProGlyIleAlaGlnGlnTrpIleGlnSer 442
        ||   |||||||||||||||||||||||||||||||||||||||||||||
    368 GAAGNTCAGAACGTCTGCAGCCGGGTATAGCCCAGCAGTGGATCCAGAGC 319

    443 LysArgGluAspIleValAsnGlnMetThrGluAlaCysLeuAsnGln.S 459
        |||||||||   |||||||||||||||||||||||||||||||||||| |
    318 AAAAGGGAAGNCATTGTGAACCAAATGACAGAAGCCTGCCTTAACCAGGT 269

    459 erLeuAspAlaLeuLeuSerArgAspLeuIleMetLysGluAspTyrGlu 475
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    268 CGCTAGATGCCCTTCTGTCCAGGGACTTGATCATGAAAGAGGACTATGAA 219

    476 LeuValSerThrLysProThrArgThrSerLysValArgGlnLeuLeuAs 492
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    218 CTTGTTAGTACCAAGCCTACAAGGACCTCAAAAGTCAGACAATTACTAGA 169

    492 pThrThrAspIleGlnGlyGluGluPheAlaLysValIleValGlnLysL 509
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    168 CACTACTGACATCCAAGGAGAAGAATTTGCCAAAGTTATAGTACAAAAAT 119

    509 euLysAspAsnLysGlnMetGlyLeuGlnProTyrProGluIleLeuVal 525
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    118 TGAAAGATAACAAACAAATGGGTCTTCAGCCTTACCCGGAAATACTTGTG 69

    526 ValSerArgSerProSerLeuAsnLeuLeuGlnAsnLysSerMet 540
        |||||||||||||||||||||||||||||||||||||||||||||
     68 GTTTCTAGATCACCATCTTTAAATTTACTTCAAAATAAAAGCATG 24

seq_name: gb_est26:AI904799

seq_documentation_block:
LOCUS       AI904799       762 bp    mRNA            EST       30-MAR-2000
DEFINITION  IL-BT067-190199-023 BT067 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AI904799
VERSION     AI904799.1  GI:6495186
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 762)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/seq/gethtml.pl?t1=IL&t2=IL-BT067-023.html
            &t3=190199&t4=1)
```

```
            Seq primer: puc 18 forward.
FEATURES             Location/Qualifiers
     source          1..762
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="BT067"
                     /sex="female"
                     /dev_stage="Adult"
                     /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No. 196
                     ,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
BASE COUNT      232 a    157 c    137 g    208 t     28 others
ORIGIN


alignment_scores:
          Quality:  876.00             Length:    226
            Ratio:  4.424                Gaps:      0
 Percent Similarity:  87.611   Percent Identity:  80.088

alignment_block:
 US-09-445-223-1 x AI904799   ..

 Align seg 1/1  to: AI904799  from: 1  to: 762

    225 ArgLysGlnProPheGluAspValThrAsnProLeuGlnIleMetTyrSe 241
        ||||||||||||||||||||||||||||||||||||||||||||||||||
      2 AGAAAACAGCCTTTTGAAGATGTCACCAATCCTTTGCAGATAATGTATAG 51

    241 rValSerGlnGlyHisArgProValIleAsnGluGluSerLeuProTyrA 258
        |   |||   ||||||||||||   |||||||||||||||||||||||||
     52 TGNGTCACANGGNCATCGNCCTGNTATTAATGAAGAAAGTTTGCCATATG 101

    258 spIleProHisArgAlaArgMetIleSerLeuIleGluSerGlyTrpAla 274
        ||   |||||||||||||||||||||||||||||||||||||||||||||
    102 ATACACCTCACCGAGCACGTATGATCTCTCTAATAGAAAGTGGATGGGCA 151

    275 GlnAsnProAspGluArgProSerPheLeuLysCysLeuIleGluLeuGl 291
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    152 CAAAATCCAGATGAAAGACCATCTTTCTTAAAATGTTTAATAGAACTTGA 201

    291 uProValLeuArgThrPheGluGluIleThrPheLeuGluAlaValIleG 308
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    202 ACCAGTTTTGAGAACATTTGAAGAGATAACTTTTCTTGAAGCTGTTATTC 251

    308 lnLeuLysLysThrLysLeuGlnSerValSerSerAlaIleHisLeuCys 324
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    252 AGCTAAAGAAAACAAAGTTACAGAGTGTTTCAAGTGCCATTCACCTATGT 301

    325 AspLysLysLysMetGluLeuSerLeuAsnIleProValAsnHisGlyPr 341
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    302 GACAAGAAGAAAATGGAATTATCTCTGAACATACCTGTAAATCATGGTCC 351

    341 oGlnGluGluSerCysGlySerSerGlnLeuHisGluAsnSerGlySerP 358
        |||||||||||||||||||||||||||||||||||||||||||||||||
    352 ACAAGAGGAATCATGTGGATCCTCTCAGCTCCATGAAAATAGTGGGTCTT 401

    358 roGluThrSerArgSerLeuProAlaProGlnAspAsnAspPheLeuSer 374
          ||||||||||||||||||||||||||||||||||||||||||::::::
    402 CTGAAACTTCAAGGTCCCTGCCAGCTCCTCAAGACAATGATTTTTTTATN 451

    375 ArgLysAlaGlnAspCysTyrPheMetLysLeuHisHisCysProGlyAs 391
           |||::::::         |||::::::||||||   |||||||||||
    452 TANAAAAGCTCAAGACTGTTNTTTTGAACCTGCATCC.TGTCCTGGNAA 500

    391 nHisSerTrpAspSerThrIleSerGlySerGlnArgAlaAlaPheCysA 408
        |:::|||||||||   ||||||   |||      :::::::::|||   |
    501 TCGCAGTTGGGATAC.ACCATTTTGGATTTNNAAAGGGTTNATTTTGGG 549
```

```
    408 spHisLysThrThrProCysSerSerAlaIleIleAsnProLeuSerThr 424
        |||||   ||||||||||||:::      |||||||||||||||||||||
    550 ATCACAANACCACTCCATGCTNTTA.AGAATAATAAATCCCCTTTCAACT 598

    425 AlaGlyAsnSerGluArgLeuGlnProGlyIleAlaGlnGlnTrpIleGl 441
        :::|||||||||:::   :::   ||||||   :::         :::
    599 GGCGGNAACTCAAAAACGTTTCNGCCCGGGNTTACCCNCCCGNGGGTCCC 648

    441 nSerLysArgGluAspIleValAsnGln 450
         |||   ||||||         ||||||
    649 AAGCNAAAGGGAAAAANTTTGGAACCAA 676

seq_name: gb_est72:BE274455

seq_documentation_block:
LOCUS       BE274455      917 bp    mRNA            EST       13-JUL-2000
DEFINITION  601120495F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:2967359 5',
            mRNA sequence.
ACCESSION   BE274455
VERSION     BE274455.1  GI:9149397
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 917)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM71  row: m  column: 24
            High quality sequence stop: 734.
FEATURES             Location/Qualifiers
     source          1..917
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2967359"
                     /clone_lib="NIH_MGC_20"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      193 a    252 c    259 g    213 t
ORIGIN


alignment_scores:
           Quality:  852.50            Length:    269
             Ratio:  4.284               Gaps:      5
 Percent Similarity:  73.978   Percent Identity:  69.888

alignment_block:
 US-09-445-223-1 x BE274455   ..

 Align seg 1/1  to: BE274455  from: 1  to: 917

      1 MetAsnGlyGluAlaIleCysSerAlaLeuProThrIleProTyrHisLy 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    176 ATGAACGGGGAGGCCATCTGCAGCGCCCTGCCCACCATTCCCTACCACAA 225

     17 sLeuAlaAspLeuArgTyrLeuSerArgGlyAlaSerGlyThrValSerS 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    226 ACTCGCCGACCTGCGCTACCTGAGCCGCGGCGCCTCTGGCACTGTGTCGT 275

     34 erAlaArgHisAlaAspTrpArgValGlnValAlaValLysHisLeuHis 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    276 CCGCCCGCCACGCAGACTGGCGCGTCCAGGTGGCCGTGAAGCACCTGCAC 325

     51 IleHisThrProLeuLeuAspSerGluArgLysAspValLeuArgGluAl 67
        |||||||||||||||||||||
    326 ATCCACACTCCGCTGCTCGACAG........................... 348

     67 aGluIleLeuHisLysAlaArgPheSerTyrIlePheProIleLeuGlyI 84

    348 .................................................. 348

     84 leCysAsnGluProGluPheLeuGlyIleValThrGluTyrMetProAsn 100

    348 .................................................. 348

    101 GlySerLeuAsnGluLeuLeuHisArgLysThrGluTyrProAspValAl 117
                                   |||||||||||||||||||||||
    349 ...........................AAAACTGAATATCCTGATGTTGC 371

    117 aTrpProLeuArgPheArgIleLeuHisGluIleAlaLeuGlyValAsnT 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    372 TTGGCCATTGAGATTTCGCATCCTGCATGAAATTGCCCTTGGTGTAAATT 421

    134 yrLeuHisAsnMetThrProProLeuLeuHisHisAspLeuLysThrGln 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    422 ACCTGCACAATATGACTCCTCCTTTACTTCATCATGACTTGAAGACTCAG 471

    151 AsnIleLeuLeuAspAsnGluPheHisValLysIleAlaAspPheGlyLe 167
        |||||||||||||||||||||||||||||||||||||||   ||||||||
    472 AATATCTTATTGGACAATGAATTTCATGTTAAGATTGCAGA.TTTGGTTT 520

    167 uSerLysTrpArgMetMetSerLeuSerGlnSerArgSerSerLysSerA 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    521 ATCAAAGTGGCGCATGATGTCCCTCTCACAGTCACGAAGTAGCAAATCTG 570

    184 laProGluGlyGlyThrIleIleTyrMetProProGluAsnTyrGluPro 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    571 CACCAGAAGGAGGGACAATTATCTATATGCCACCTGAAAACTATGAACCT 620

    201 GlyGlnLysSerArgAlaSerIleLysHisAspIleTyrSerTyrAlaVa 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    621 GGACAAAAATCAAGGGCCAGTATCAAGCACGATATATATAGCTATGCAGT 670

    217 lIleThrTrp..GluValLeuSerArgLysGlnProPheGluAspValTh 233
        ||||||||||     |||||||||   ::::::   ||||||||||||||
    671 TATCACATGGGCACAGTGTTATCC...GAAAACAGCTTTGAAGATGTCAC 717

    233 rAsnProLeuGlnIleMetTyrSerValSerGlnGlyHisArgProValI 250
        |:::||||||||||||      :::|||   :::            ::::
    718 CCATCCTTTGCAGATATGTTAGGGTGTCCCCAGG......CTCGACCTGT 761

    250 leAsnGluGluSerLeuProTyrAspIleProHisArgAlaArgMetIle 266
        ::      |||:::||||||      :::   |||:::|||      :::
    762 TTTATGGTGAAGGTTTGCCTATG...TTTTCTCACGAGGCCGGTTGTCTC 808

    267 SerLeu 268
        ||||||
    809 TCATTG 814

seq_name: gb_est41:AW085560

seq_documentation_block:
LOCUS       AW085560      606 bp    mRNA            EST       09-MAR-2000
DEFINITION  wy67c04.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
            IMAGE:2553606 3' similar to TR:O43353 O43353 SERINE/THREONINE
            KINASE RICK. ;, mRNA sequence.
ACCESSION   AW085560
```

```
VERSION     AW085560.1  GI:6040712
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 606)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 885   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 470.
FEATURES             Location/Qualifiers
     source          1..606
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2553606"
                     /clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from five normalized
                     libraries were mixed, and ss circles were made in vitro.
                     Following HAP purification, this DNA was used as tracer in
                     a subtractive hybridization reaction. The driver was
                     PCR-amplified cDNAs from pools of 5,000 clones made from
                     the same 5 libraries. The pools consisted of the following
                     libraries and cloneIDs: Soares NbHSF pool 1:
                     309384-310919, 323208-325895 Soares Nb2HP pool 1:
                     145032-147335, 147720-148103, 148872-149255, 15002 -
                     150407, 151176-152327 Soares Nb2HF8-9W pool 1:
                     758280-760583, 772104-774407 Soares NbHPA pool 1:
                     304776-306311, 320136-322823, 326280-326663 Soares NbHOT
                     pool 1:  723720-726407, 739080-740999 Subtraction by Bento
                     Soares and M. Fatima Bonaldo."
BASE COUNT      166 a    114 c    122 g    203 t      1 others
ORIGIN


alignment_scores:
          Quality:  838.00             Length:    166
            Ratio:  5.110                Gaps:      1
 Percent Similarity:  98.795   Percent Identity: 98.795

alignment_block:
 US-09-445-223-1 x AW085560/rev   ..

 Align seg 1/1  to reverse of: AW085560  from: 1  to: 606

    376 LysAlaGlnAspCysTyrPheMetLysLeuHisHisCysProGlyAsnHi 392
        ||||||||||||||||||   |||||||||||||||||||||||||||||
    606 AAAGCTCAAGACTGTTATTNTATGAAGCTGCATCACTGTCCTGGAAATCA 557

    392 sSerTrpAspSerThrIleSerGlySerGlnArgAlaAlaPheCysAspH 409
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    556 CAGTTGGGATAGCACCATTTCTGGATCTCAAAGGGCTGCATTCTGTGATC 507

    409 isLysThrThrProCysSerSerAlaIleIleAsnProLeuSerThrAla 425
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    506 ACAAGACCACTCCATGCTCTTCAGCAATAATAAATCCACTCTCAACTGCA 457

    426 GlyAsnSerGluArgLeuGlnProGlyIleAlaGlnGlnTrpIleGlnSe 442
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    456 GGAAACTCAGAACGTCTGCAGCCTGGTATAGCCCAGCAGTGGATCCAGAG 407

    442 rLysArgGluAspIleValAsnGlnMetThrGluAlaCysLeuAsnGlnS 459
```

```
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    406 CAAAAGGGAAGACATTGTGAACCAAATGACAGAAGCCTGCCTTAACCAGT 357

    459 erLeuAspAlaLeuLeuSerArgAsp.LeuIleMetLysGluAspTyrGl 475
        |||||||||||||||||||||||||| |||||||||||||||||||||||
    356 CGCTAGATGCCCTTCTGTCCAGGGACTTTGATCATGAAAGAGGACTATGA 307

    475 uLeuValSerThrLysProThrArgThrSerLysValArgGlnLeuLeuA 492
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    306 ACTTGTTAGTACCAAGCCTACAAGGACCTCAAAAGTCAGACAATTACTAG 257

    492 spThrThrAspIleGlnGlyGluGluPheAlaLysValIleValGlnLys 508
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    256 ACACTACTGACATCCAAGGAGAAGAATTTGCCAAAGTTATAGTACAAAAA 207

    509 LeuLysAspAsnLysGlnMetGlyLeuGlnProTyrProGluIleLeuVa 525
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    206 TTGAAAGATAACAAACAAATGGGTCTTCAGCCTTACCCGGAAATACTTGT 157

    525 lValSerArgSerProSerLeuAsnLeuLeuGlnAsnLysSerMet 540
        ||||||||||||||||||||||||||||||||||||||||||||||
    156 GGTTTCTAGATCACCATCTTTAAATTTACTTCAAAATAAAAGCATG 111

seq_name: gb_est5:AA315575

seq_documentation_block:
LOCUS       AA315575      484 bp    mRNA            EST       19-APR-1997
DEFINITION  EST187344 Colon carcinoma (HCC) cell line II Homo sapiens cDNA 5'
            end, mRNA sequence.
ACCESSION   AA315575
VERSION     AA315575.1  GI:1967904
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 484)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
            ,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
            ,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
            Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
            ,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A.,
            Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M.,
            Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W.
            , Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
            Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L.,
            Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon
            ,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
            Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   96026280
COMMENT     Other_ESTs: THC188294
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1..484
                     /organism="Homo sapiens"
```

```
                     /db_xref="ATCC (inhost):110668"
                     /db_xref="taxon:9606"
                     /clone_lib="Colon carcinoma (HCC) cell line II"
                     /tissue_type="colon"
                     /cell_type="KM12C"
                     /cell_line="KM12C(HCC)-parental human colon carcinoma
                     ;Dukes B2"
                     /note="Organ: colon; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI"
BASE COUNT      136 a    109 c    101 g    136 t       2 others
ORIGIN

alignment_scores:
           Quality:  820.00             Length:    161
             Ratio:  5.157                Gaps:      1
 Percent Similarity:  98.758   Percent Identity: 98.137

alignment_block:
 US-09-445-223-1 x AA315575   ..

 Align seg 1/1  to: AA315575  from: 1  to: 484

      27 GlyAlaSerGlyThrValSerSerAlaArgHisAlaAspTrpArgValGl 43
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       2 GGCGCCTCTGGCACTGTGTCGTCCGCCCGCCACGCAGACTGGCGCGTCCA 51

      43 nValAlaValLysHisLeuHisIleHisThrProLeuLeuAspSerGluA 60
         ||||||||||||||||||||||||||||||||||||||||||||||   |
      52 GGTGGCCGTGAAGCACCTGCACATCCACACTCCGCTGCTCGACAGTNAAA 101

      60 rgLysAspValLeuArgGluAlaGluIleLeuHisLysAlaArgPheSer 76
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     102 GAAAGGATGTCTTAAGAGAAGCTGAAATTTTACACAAAGCTAGATTTAGT 151

      77 TyrIlePheProIleLeuGlyIleCysAsnGluProGluPheLeuGlyIl 93
         ||||||:::|||||||||||||||||||||||||||||||||||||||||
     152 TACATTCTTCCAATTTTGGGAATTTGCAATGAGCCTGAATTTTTGGGAAT 201

      93 eValThrGluTyrMetProAsnGlySerLeuAsnGluLeuLeuHisArgL 110
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     202 AGTTACTGAATACATGCCAAATGGNTCATTAAATGAACTCCTACATAGGA 251

     110 ysThrGluTyrProAspValAlaTrpProLeuArgPheArgIleLeuHis 126
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     252 AAACTGAATATCCTGATGTTGCTTGGCCATTGAGATTTCGCATCCTGCAT 301

     127 GluIleAlaLeuGlyValAsnTyrLeuHisAsnMetThrProProLeuLe 143
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     302 GAAATTGCCCTTGGTGTAAATTACCTGCACAATATGACTCCTCCTTTACT 351

     143 uHisHisAspLeuLysThrGlnAsnIleLeuLeuAspAsnGluPheHisV 160
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     352 TCATCATGACTTGAAGACTCAGAATATCTTATTGGACAATGAATTTCATG 401

     160 alLysIleAlaAspPheGlyLeuSerLysTrpArgMetMetSerLeuSer 176
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     402 TTAAGATTGCAGATTTTGGTTTATCAAAGTGGCGCATGATGTCCCTCTCA 451

     177 GlnSerArgSerSer.LysSerAlaProGlu 186
         ||||||||||||||| |||||||||||||||
     452 CAGTCACGAAGTAGCAAAATCTGCACCAGAA 482

seq_name: gb_est10:AA655189

seq_documentation_block:
LOCUS       AA655189        503 bp    mRNA            EST       04-NOV-1997
DEFINITION  vv13a12.r1 Stratagene mouse heart (#937316) Mus musculus cDNA clone
            IMAGE:1211518 5' similar to TR:G1236943 G1236943 RIP PROTEIN
            KINASE. ;, mRNA sequence.
ACCESSION   AA655189
VERSION     AA655189.1  GI:2591343
KEYWORDS    EST.
```

```
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 503)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:647862
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 436.
FEATURES             Location/Qualifiers
     source          1. .503
                     /organism="Mus musculus"
                     /strain="NIH/Swiss"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1211518"
                     /clone_lib="Stratagene mouse heart (#937316)"
                     /sex="pooled"
                     /tissue_type="heart"
                     /dev_stage="13 day embryos"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: heart; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. 93 pooled NIH/Swiss 13 day embryo hearts.
                     Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTTTT 3'"
BASE COUNT      158 a    102 c    105 g    138 t
ORIGIN


alignment_scores:
           Quality:  811.00             Length:    167
             Ratio:  5.006                Gaps:      0
 Percent Similarity:  97.006   Percent Identity: 91.617

alignment_block:
 US-09-445-223-1 x AA655189   ..

 Align seg 1/1  to: AA655189  from: 1  to: 503

     142 LeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeuAspAsnGluPh 158
         ||||||||||||:::|||||||||||||||||||||||||||||||||||
       2 CTGCTTCATCATGAACTAAAGACTCAGAATATCTTGTTGGATAATGAATT 51

     158 eHisValLysIleAlaAspPheGlyLeuSerLysTrpArgMetMetSerL 175
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      52 TCATGTTAAGATTGCAGATTTTGGTTTATCAAAATGGCGCATGATGTCAC 101

     175 euSerGlnSerArgSerSerLysSerAlaProGluGlyGlyThrIleIle 191
         |||||||||||||||||   ||||||||||||||||||||||||||||||
     102 TCTCTCAATCACGAAGTTACAAATCTGCACCCGAAGGAGGAACAATCATC 151

     192 TyrMetProProGluAsnTyrGluProGlyGlnLysSerArgAlaSerIl 208
         ||||||||||||||||||||||||||||||||||||||||||||||||::
     152 TATATGCCACCTGAGAACTATGAGCCAGGACAGAAATCAAGGGCCAGTGT 201

     208 eLysHisAspIleTyrSerTyrAlaValIleThrTrpGluValLeuSerA 225
         :||||||||||||||||||||||||||||||   ||||||||||||||||
```

```
    202 GAAGCATGATATATACAGCTATGCAGTTATCATGTGGGAAGTGTTATCCA 251

    225 rgLysGlnProPheGluAspValThrAsnProLeuGlnIleMetTyrSer 241
        |||||||||||||||||:::||||||||||||||||||||||||||||||
    252 GAAAACAGCCTTTTGAAGAGGTCACCAATCCTTTGCAGATCATGTACAGT 301

    242 ValSerGlnGlyHisArgProValIleAsnGluGluSerLeuProTyrAs 258
        |||||||||||||||||||||      :::||||||:::||||||:::||
    302 GTGTCCCAGGGACATCGACCTGACACCAGTGAGGAGAATTTGCCATTTGA 351

    258 pIleProHisArgAlaArgMetIleSerLeuIleGluSerGlyTrpAlaG 275
        |||||||||||||:::   |||||||||||||||:::|||||||||||||
    352 TATACCTCATCGAGGTCTCATGATCTCTCTAATACAGAGTGGATGGGCGC 401

    275 lnAsnProAspGluArgProSerPheLeuLysCysLeuIleGluLeuGlu 291
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    402 AAAACCCAGATGAAAGACCATCCTTTTTGAAATGCTTAATAGAACTTGAA 451

    292 ProValLeuArgThrPheGluGluIleThrPheLeuGluAlaValIleGl 308
        |||||||||||||||||||||:::||||||||||||||||||||||||||
    452 CCAGTTCTGAGAACATTTGAAGACATAACTTTCCTTGAGGCTGTTATTCA 501

    308 n 308
        |
    502 G 502

seq_name: gb_est45:AW389863

seq_documentation_block:
LOCUS       AW389863       510 bp    mRNA            EST       04-FEB-2000
DEFINITION  RC2-ST0173-041099-011-e06 ST0173 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW389863
VERSION     AW389863.1  GI:6894522
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 510)
  AUTHORS   HCGP http://www.ludwig.org.br/ORESTES.
  TITLE     The FAPESP/LICR Human Cancer Genome Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC2&t2=RC2-ST0173-
            041099-011-e06&t3=1999-10-04&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 13
            High quality sequence stop: 510.
FEATURES             Location/Qualifiers
     source          1..510
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="ST0173"
                     /dev_stage="Adult"
                     /note="Organ: stomach; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT      127 a    108 c    128 g    147 t
ORIGIN
```

```
alignment_scores:
           Quality:  798.50              Length:    221
             Ratio:  4.810                 Gaps:      2
 Percent Similarity:  75.113    Percent Identity:  72.851

alignment_block:
 US-09-445-223-1 x AW389863/rev   ..

 Align seg 1/1  to reverse of: AW389863  from: 1  to: 510

     19 AlaAspLeuArgTyrLeuSerArgGlyAlaSerGlyThrValSerSerAl 35
        |||:::|||||||||:::|||:::||||||||||||||||||||||||||
    509 GCCGAACTGCGCTACATGAGCCACGGCGCCTCAGGCACTGTGTCGTCCGC 460

     35 aArgHisAlaAspTrpArgValGlnValAlaValLysHisLeuHisIleH 52
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    459 CCGCCACGCAGATTGGCGCGTCCAGGTGGCCGTGAAGCACCTGCACATCC 410

     52 isThrProLeuLeuAspSerGluArgLysAspValLeuArgGluAlaGlu 68
        |||||||||||||||||
    409 ACACTCCGCTGCTCGACAG............................... 391

     69 IleLeuHisLysAlaArgPheSerTyrIlePheProIleLeuGlyIleCy 85

    391 .................................................. 391

     85 sAsnGluProGluPheLeuGlyIleValThrGluTyrMetProAsnGlyS 102

    391 .................................................. 391

    102 erLeuAsnGluLeuLeuHisArgLysThrGluTyrProAspValAlaTrp 118
                               |||||||||||||||||||||||||||
    390 .......................AAAACTGAATATCCTGATGTTGCTTGG 364

    119 ProLeuArgPheArgIleLeuHisGluIleAlaLeuGlyValAsnTyrLe 135
        |||||||||||||||||||||:::|||||||||:::||||||||||||||
    363 CCATTGAGATTTCGCATCCTGAATGAAATTGCCATTGGTGTAAATTACCT 314

    135 uHisAsnMetThrProProLeuLeuHisHisAspLeuLysThrGln.Asn 151
        ||||||||||||||||||||||||||||||||||   ||||||||| |||
    313 GCACAATATGACTCCTCCTTTACTTCATCATGACTGGAAGACTCAAGAAT 264

    152 IleLeuLeuAspAsnGluPheHisValLysIleAlaAspPheGlyLeuSe 168
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    263 ATCTTATTGGACAATGAATTTCATGTTAAGATTGCAGATTTTGGTTTATC 214

    168 rLysTrpArgMetMetSerLeuSerGlnSerArgSerSerLysSerAlaP 185
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    213 AAAGTGGCGCATGATGTCCCTCTCACAGTCACGAAGTAGCAAATCTGCAC 164

    185 roGluGlyGlyThrIleIleTyrMetProProGluAsnTyrGluProGly 201
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    163 CAGAAGGAGGGACAATTATCTATATGCCACCTGAAAACTATGAACCTGGA 114

    202 GlnLysSerArgAlaSerIleLysHisAspIleTyrSerTyrAlaValIl 218
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    113 CAAAAATCCAGGGCCAGTATCAAGCACGATATATATAGCTATGCAGTTAT 64

    218 eThrTrpGluValLeuSerArgLysGlnProPheGluAspValThrAsnP 235
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     63 CACATGGGAAGTGCTATCCAGAAAACAGCCTTTTGAAGATGTCACCAATC 14

    235 roLeuGlnIle 238
        ||   ||||||
     13 CTTGCCAGATA 3
```

```
OM of: US-09-445-223-1 to: Issued_Patents_NA:*   out_format : pfs

Date: Jun 13, 2001  6:36 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn1_1/USPTO_spool/US09445223/runat_13062001_092033_29663/app_query.fasta_1.604
 -DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni
 -GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000
 -LOOPEXT=0.000 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000
 -XGAPEXT=0.500 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000
 -YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1
 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200
 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
 -OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=2000000000
 -USER=US09445223_@CGN1_1_88 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
 -WAIT -THREADS=1

Search information block:
Query: US-09-445-223-1
Query length: 540
Database: Issued_Patents_NA:*
Database sequences: 302621
Database length: 87301344
Search time (sec): 68.920000


score_list:
Sequence             Strd Orig    ZScore   EScore Len  ! Documentation  ..
/cgn1_7/ptodata/1/ina/6A_COMB.seq:US-09-019-942-2 + 2817.00 5441.25  1.7e-295  1931  !
/cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-188-930-257 +  579.00 1078.43  1.7e-52   3516  !
/cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-188-930-66 +  548.00 1026.11  1.4e-49   1888  !
/cgn1_7/ptodata/1/ina/5A_COMB.seq:US-08-444-005-14 +  397.00  729.91  4.5e-33   2268  !
/cgn1_7/ptodata/1/ina/6A_COMB.seq:US-09-329-418-2 +  395.00  730.86  4.0e-33   1557  !
/cgn1_7/ptodata/1/ina/6A_COMB.seq:US-09-329-418-1 +  392.50  723.62  1.0e-32   1873  !
/cgn1_7/ptodata/1/ina/5A_COMB.seq:US-07-928-464-1 +  376.00  685.92  1.3e-30   2890  !
/cgn1_7/ptodata/1/ina/PCTUS_COMB.seq:PCT-US93-07347-1 +  376.00  685.92  1.3e-30   2890
/cgn1_7/ptodata/1/ina/5A_COMB.seq:US-08-003-311B-1 +  376.00  685.30  1.4e-30   3033  !
/cgn1_7/ptodata/1/ina/5A_COMB.seq:US-08-261-432-1 +  376.00  685.30  1.4e-30   3033  !
/cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-132-118-1 +  369.00  676.94  4.0e-30   2016  !
/cgn1_7/ptodata/1/ina/5A_COMB.seq:US-08-444-005-16 +  369.00  676.19  4.4e-30   2137  !
/cgn1_7/ptodata/1/ina/6A_COMB.seq:US-09-161-443-1 +  369.00  673.58  6.2e-30   2617  !
/cgn1_7/ptodata/1/ina/5A_COMB.seq:US-08-700-575-31 +  367.00  702.19  1.6e-31   210  !
/cgn1_7/ptodata/1/ina/5B_COMB.seq:US-08-685-625A-1 +  315.50  570.35  3.5e-24   2443  !
/cgn1_7/ptodata/1/ina/5B_COMB.seq:US-08-685-625A-5 +  312.00  562.46  9.6e-24   2656  !
/cgn1_7/ptodata/1/ina/5A_COMB.seq:US-08-395-580-1 +  302.00  539.86  1.7e-22   3389  !
/cgn1_7/ptodata/1/ina/5A_COMB.seq:US-08-205-018-1 +  302.00  539.73  1.8e-22   3426  !
/cgn1_7/ptodata/1/ina/6A_COMB.seq:US-09-221-235-6 +  291.50  531.15  5.3e-22   1365  !
/cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-221-928-6 +  291.50  531.15  5.3e-22   1365  !
/cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-221-527-6 +  291.50  531.15  5.3e-22   1365  !
/cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-221-236-6 +  291.50  531.15  5.3e-22   1365  !
/cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-221-416-6 +  291.50  531.15  5.3e-22   1365  !
/cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-221-245-6 +  291.50  531.15  5.3e-22   1365  !
/cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-163-115-6 +  291.50  531.15  5.3e-22   1365  !
/cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-221-528-6 +  291.50  531.15  5.3e-22   1365  !
/cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-593-553-6 +  291.50  531.15  5.3e-22   1365  !
/cgn1_7/ptodata/1/ina/6A_COMB.seq:US-09-221-235-4 +  291.50  525.48  1.1e-21   2120  !
/cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-221-928-4 +  291.50  525.48  1.1e-21   2120  !
/cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-221-527-4 +  291.50  525.48  1.1e-21   2120  !
/cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-221-236-4 +  291.50  525.48  1.1e-21   2120  !
/cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-221-416-4 +  291.50  525.48  1.1e-21   2120  !
/cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-221-245-4 +  291.50  525.48  1.1e-21   2120  !
/cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-163-115-4 +  291.50  525.48  1.1e-21   2120  !
/cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-221-528-4 +  291.50  525.48  1.1e-21   2120  !
/cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-593-553-4 +  291.50  525.48  1.1e-21   2120  !
/cgn1_7/ptodata/1/ina/5A_COMB.seq:US-08-587-889-1 +  283.50  503.12  1.9e-20   3590  !
/cgn1_7/ptodata/1/ina/PCTUS_COMB.seq:PCT-US96-09193-1 +  283.50  503.12  1.9e-20   3590
/cgn1_7/ptodata/1/ina/5B_COMB.seq:US-08-888-982A-42 +  280.50  501.90  2.3e-20   2510  !
/cgn1_7/ptodata/1/ina/5A_COMB.seq:US-08-276-151-1 +  271.50  482.18  2.8e-19   2977  !
/cgn1_7/ptodata/1/ina/5A_COMB.seq:US-08-306-691B-17 +  271.50  482.18  2.8e-19   2977
/cgn1_7/ptodata/1/ina/5B_COMB.seq:US-08-756-806A-64 +  271.50  482.18  2.8e-19   2977
/cgn1_7/ptodata/1/ina/6A_COMB.seq:US-08-328-239A-4 +  271.50  482.18  2.8e-19   2977
/cgn1_7/ptodata/1/ina/6A_COMB.seq:US-09-143-214-64 +  271.50  482.18  2.8e-19   2977
/cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-209-668-12 +  271.50  482.18  2.8e-19   2977

seq_name: /cgn1_7/ptodata/1/ina/6A_COMB.seq:US-09-019-942-2

seq_documentation_block:
; Sequence 2, Application US/09019942
; Patent No. 6033855
;  GENERAL INFORMATION:
;    APPLICANT:  Bertin, John
;    TITLE OF INVENTION:  GENES ENCODING CASPASE RECRUITMENT
;    TITLE OF INVENTION:  DOMAIN POLYPEPTIDES
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/019,942
;      FILING DATE:  06-FEB-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meiklejohn, Ph.D., Anita L.
;      REGISTRATION NUMBER:  35,283
;      REFERENCE/DOCKET NUMBER:  07334/068001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1931 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-019-942-2


alignment_scores:
           Quality:  2817.00             Length:     540
             Ratio:  5.226                 Gaps:       0
 Percent Similarity:  99.815   Percent Identity: 99.630

alignment_block:
 US-09-445-223-1 x US-09-019-942-2   ..

 Align seg 1/1  to: US-09-019-942-2  from: 1  to: 1931

       1 MetAsnGlyGluAlaIleCysSerAlaLeuProThrIleProTyrHisLy 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     214 ATGAACGGGGAGGCCATCTGCAGCGCCCTGCCCACCATTCCCTACCACAA 263
            \
      17 sLeuAlaAspLeuArgTyrLeuSerArgGlyAlaSerGlyThrValSerS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     264 ACTCGCCGACCTGCGCTACCTGAGCCGCGGCGCCTCTGGCACTGTGTCGT 313

      34 erAlaArgHisAlaAspTrpArgValGlnValAlaValLysHisLeuHis 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     314 CCGCCCGCCACGCAGACTGGCGCGTCCAGGTGGCCGTGAAGCACCTGCAC 363
```

```
   51 IleHisThrProLeuLeuAspSerGluArgLysAspValLeuArgGluAl 67
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  364 ATCCACACTCCGCTGCTCGACAGTGAAAGAAAGGATGTCTTAAGAGAAGC 413

   67 aGluIleLeuHisLysAlaArgPheSerTyrIlePheProIleLeuGlyI 84
      ||||||||||||||||||||||||||||||||||:::|||||||||||||
  414 TGAAATTTTACACAAAGCTAGATTTAGTTACATTCTTCCAATTTTGGGAA 463

   84 leCysAsnGluProGluPheLeuGlyIleValThrGluTyrMetProAsn 100
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  464 TTTGCAATGAGCCTGAATTTTTGGGAATAGTTACTGAATACATGCCAAAT 513

  101 GlySerLeuAsnGluLeuLeuHisArgLysThrGluTyrProAspValAl 117
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  514 GGATCATTAAATGAACTCCTACATAGGAAAACTGAATATCCTGATGTTGC 563

  117 aTrpProLeuArgPheArgIleLeuHisGluIleAlaLeuGlyValAsnT 134
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  564 TTGGCCATTGAGATTTCGCATCCTGCATGAAATTGCCCTTGGTGTAAATT 613

  134 yrLeuHisAsnMetThrProProLeuLeuHisHisAspLeuLysThrGln 150
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  614 ACCTGCACAATATGACTCCTCCTTTACTTCATCATGACTTGAAGACTCAG 663

  151 AsnIleLeuLeuAspAsnGluPheHisValLysIleAlaAspPheGlyLe 167
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  664 AATATCTTATTGGACAATGAATTTCATGTTAAGATTGCAGATTTTGGTTT 713

  167 uSerLysTrpArgMetMetSerLeuSerGlnSerArgSerSerLysSerA 184
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  714 ATCAAAGTGGCGCATGATGTCCCTCTCACAGTCACGAAGTAGCAAATCTG 763

  184 laProGluGlyGlyThrIleIleTyrMetProProGluAsnTyrGluPro 200
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  764 CACCAGAAGGAGGGACAATTATCTATATGCCACCTGAAAACTATGAACCT 813

  201 GlyGlnLysSerArgAlaSerIleLysHisAspIleTyrSerTyrAlaVa 217
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  814 GGACAAAAATCAAGGGCCAGTATCAAGCACGATATATATAGCTATGCAGT 863

  217 lIleThrTrpGluValLeuSerArgLysGlnProPheGluAspValThrA 234
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  864 TATCACATGGGAAGTGTTATCCAGAAAACAGCCTTTTGAAGATGTCACCA 913

  234 snProLeuGlnIleMetTyrSerValSerGlnGlyHisArgProValIle 250
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  914 ATCCTTTGCAGATAATGTATAGTGTGTCACAAGGACATCGACCTGTTATT 963

  251 AsnGluGluSerLeuProTyrAspIleProHisArgAlaArgMetIleSe 267
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  964 AATGAAGAAAGTTTGCCATATGATATACCTCACCGAGCACGTATGATCTC 1013

  267 rLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerPheL 284
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1014 TCTAATAGAAAGTGGATGGGCACAAAATCCAGATGAAAGACCATCTTTCT 1063

  284 euLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGluIle 300
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1064 TAAAATGTTTAATAGAACTTGAACCAGTTTTGAGAACATTTGAAGAGATA 1113

  301 ThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSerVa 317
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1114 ACTTTTCTTGAAGCTGTTATTCAGCTAAAGAAAACAAAGTTACAGAGTGT 1163

  317 lSerSerAlaIleHisLeuCysAspLysLysLysMetGluLeuSerLeuA 334
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1164 TTCAAGTGCCATTCACCTATGTGACAAGAAGAAAATGGAATTATCTCTGA 1213

  334 snIleProValAsnHisGlyProGlnGluGluSerCysGlySerSerGln 350
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1214 ACATACCTGTAAATCATGGTCCACAAGAGGAATCATGTGGATCCTCTCAG 1263

  351 LeuHisGluAsnSerGlySerProGluThrSerArgSerLeuProAlaPr 367
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1264 CTCCATGAAAATAGTGGTTCTCCTGAAACTTCAAGGTCCCTGCCAGCTCC 1313

  367 oGlnAspAsnAspPheLeuSerArgLysAlaGlnAspCysTyrPheMetL 384
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1314 TCAAGACAATGATTTTTTATCTAGAAAAGCTCAAGACTGTTATTTTATGA 1363

  384 ysLeuHisHisCysProGlyAsnHisSerTrpAspSerThrIleSerGly 400
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1364 AGCTGCATCACTGTCCTGGAAATCACAGTTGGGATAGCACCATTTCTGGA 1413

  401 SerGlnArgAlaAlaPheCysAspHisLysThrThrProCysSerSerAl 417
      |||||||||||||||||||||||||||||||||   |||||·||||||||
 1414 TCTCAAAGGGCTGCATTCTGTGATCACAAGACCATTCCATGCTCTTCAGC 1463

  417 aIleIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeuGlnProG 434
      ||||||||||||||||||||||||||||||||||||||·|||||||||||
 1464 AATAATAAATCCACTCTCAACTGCAGGAAACTCAGAACGTCTGCAGCCTG 1513

  434 lyIleAlaGlnGlnTrpIleGlnSerLysArgGluAspIleValAsnGln 450
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1514 GTATAGCCCAGCAGTGGATCCAGAGCAAAAGGGAAGACATTGTGAACCAA 1563

  451 MetThrGluAlaCysLeuAsnGlnSerLeuAspAlaLeuLeuSerArgAs 467
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1564 ATGACAGAAGCCTGCCTTAACCAGTCGCTAGATGCCCTTCTGTCCAGGGA 1613

  467 pLeuIleMetLysGluAspTyrGluLeuValSerThrLysProThrArgT 484
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1614 CTTGATCATGAAAGAGGACTATGAACTTGTTAGTACCAAGCCTACAAGGA 1663

  484 hrSerLysValArgGlnLeuLeuAspThrThrAspIleGlnGlyGluGlu 500
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1664 CCTCAAAAGTCAGACAATTACTAGACACTACTGACATCCAAGGAGAAGAA 1713

  501 PheAlaLysValIleValGlnLysLeuLysAspAsnLysGlnMetGlyLe 517
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1714 TTTGCCAAAGTTATAGTACAAAAATTGAAAGATAACAAACAAATGGGTCT 1763

  517 uGlnProTyrProGluIleLeuValValSerArgSerProSerLeuAsnL 534
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1764 TCAGCCTTACCCGGAAATACTTGTGGTTTCTAGATCACCATCTTTAAATT 1813

  534 euLeuGlnAsnLysSerMet 540
      ||||||||||||||||||||
 1814 TACTTCAAAATAAAAGCATG 1833

seq_name: /cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-188-930-257

seq_documentation_block:
; Sequence 257, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
;  APPLICANT: Watson, James D.
;  APPLICANT: Strachan, Lorna
;  APPLICANT: Sleeman, Matthew
;  APPLICANT: Onrust, Rene
;  APPLICANT: Murison, James Greg
;  TITLE OF INVENTION: Compositions Isolated From Skin Cells
;  TITLE OF INVENTION: and Methods For Their Use
;  FILE REFERENCE: 11000.1011c1
;  CURRENT APPLICATION NUMBER: US/09/188,930A
;  CURRENT FILING DATE: 1998-11-09
;  NUMBER OF SEQ ID NOS: 348
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 257
;   LENGTH: 3516
;   TYPE: DNA
;   ORGANISM: Mouse
US-09-188-930-257
```

```
alignment_scores:
           Quality:  579.00           Length:    540
             Ratio:  1.969              Gaps:     26
 Percent Similarity:  54.444  Percent Identity: 34.074

alignment_block:
 US-09-445-223-1 x US-09-188-930-257

 Align seg 1/1  to: US-09-188-930-257  from: 1  to: 3516

 10 LeuProThrIleProTyrHisLysLeuAlaAspLeuArgTyrLeuSerAr 26
    |||   |||:::         ::::::|||      :::   ::::::
 43 CTGCGCACCTTCGACGCCGGCGAATTCGCAGGCTGGGAGAAGGTGGGCTC 92

 26 gGlyAlaSerGlyThrValSerSerAlaArgHisAlaAspTrpArgValG 43
     |||:::   |||   |||   ::::::||||||:::   |||::::::
 93 GGGCGGCTTCGGGCAGGTGTACAAGGTGCGCCATGTGCACTGGAAGACGT 142

 43 lnValAlaValLys.........HisLeuHisIleHisThrProLeuLeu 56
      :::|||:::|||            ||||||:::
143 GGCTCGCGATCAAGTGCTCGCCCAGTCTGCACGTCGAC............ 180

 57 AspSerGluArgLysAspValLeuArgGluAlaGluIleLeuHisLysAl 73
    |||   ||||||   ::::::|||:::||||||:::   ::::::   ||
181 GACAGGGAACGAATGGAGCTCCTGGAGGAAGCTAAGAAGATGGAGATGGC 230

 73 aArgPheSerTyrIlePheProIleLeuGlyIleCysAsnGluProGluP 90
    |:::|||   ||||||:::|||:::   |||||||||:::||||||
231 CAAGTTCCGATACATTCTACCTGTGTACGGCATATGCCAGGAACCT.... 276

 90 heLeuGlyIleValThrGluTyrMetProAsnGlySerLeuAsnGluLeu 106
      :::|||:::|||   |||||||||   :::|||||||||::::::|||
277 ..GTCGGCTTGGTCATGGAGTACATGGAGACAGGCTCCCTGGAGAAGCTG 324

107 LeuHisArgLysThrGluTyrProAspValAlaTrpProLeuArgPheAr 123
    |||      :::               :::   |||   |||||||||||
325 CTGGCCTCAGAG............CCATTGCCTTGGGACCTGCGCTTTCG 362

123 gIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisAsnMetThrP 140
    ||||:::||||||   |||:::|||:::|||:::||||||   |||:::|
363 CATCGTGCACGAGACAGCCGTGGGCATGAACTTCCTGCATTGCATGTCTC 412

140 roProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeuAspAsn 156
    ||||||||||||||   |||||||||      |||||||||||||||
413 CGCCACTGCTGCACCTAGACCTGAAGCCAGCGAACATCCTGCTGGATGCC 462

157 GluPheHisValLysIleAlaAspPheGlyLeuSerLysTrpArgMetMe 173
    ::::::||||||||||||:::||||||||||||:::|||   :::   ||
463 CACTACCATGTCAAGATTTCTGACTTTGGGCTGGCCAAGTGCAATGGCAT 512

173 tSerLeuSerGlnSerArgSerSerLysSerAlaProGluGly......G 188
    ||||   |||::::::   |||               :::|||      |
513 GTCCCACTCTCATGACCTCAGCATG............GATGGCCTGTTTG 550

188 lyThrIleIleTyrMetProProGluAsnTyrGluProGlyGlnLysSer 204
    ||||||||   |||:::|||||||||:::   :::      :::||||||
551 GTACAATCGCTTACCTCCCTCCAGAGCGAATTCGT......GAGAAGAGC 594

205 ArgAla...SerIleLysHisAspIleTyrSerTyrAlaValIleThrTr 220
    |||      :::   |||||||||:::||||||:::|||::::::   ||
595 CGCTTGTTTGACACCAAACATGATGTATACAGCTTCGCCATTGTGATCTG 644

220 pGluValLeuSerArgLysGlnProPheGluAspValThrAsnProLeuG 237
    |   ||||||::::::|||:::||||||   |||      |||   |||:
645 GGGTGTGCTTACACAGAAGAAGCCATTTGCAGATGAAAAGAACATCCTAC 694

237 lnIleMetTyrSerValSerGlnGlyHisArgProValIleAsnGluGlu 253
    ::||||||   :::|||   :::||||||||||||            |||
695 ACATCATGATGAAAGTGGTAAAGGGCCACCGCCCA............GAG 732

254 SerLeuProTyrAspIleProHisArg.........AlaArgMetIleSe 267
          |||         |||:::            |||   :::|||::
733 CTGCCACCCATCTGCAGACCCCGGCCGCGTGCCTGTGCCAGCCTGATAGG 782

 267 rLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerPheL 284
     :|||::::::      |||      :::|||:::   ||||||
 783 GCTCATGCAACGGTGCTGGCATGCAGACCCACAGGTGCGGCCC....... 825

 284 euLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGluIle 300
                                        ||||||:::||||||
 826 ...................................ACCTTCCAAGAAATT 840

 301 ThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSerVa 317
     |||
 841 ACC............................................... 843

 317 lSerSerAlaIleHisLeuCysAspLysLysLysMetGluLeuSerLeuA 334
      |||::::::      ||||||:::|||         |||::::::
 844 .TCTGAAACAGAAGACCTTTGTGAGAAGCCTGATGAGGAGGTGAAA.... 888

 334 snIleProValAsnHisGlyProGlnGluGluSerCysGlySerSerGln 350
             :::   |||   |||   |||:::|||      :::|||:::
 889 .....GACCTGGCTCATGAGCCAGGCGAGAAAAGCTCTCTAGAGTCCAAG 933

 351 LeuHisGluAsnSerGlySerProGluThrSerArgSerLeuProAlaPr 367
        :::   :::      |||   :::   :::|||:::         ||
 934 AGTGAGGCCAGGCCCGAGTCCTCACGCCTCAAGCGCGCCTCTGCTCCCCC 983

 367 oGlnAspAsnAsp.PheLeuSerArgLysAlaGlnAspCysTyrPheMet 383
     |   |||||||||    :::|||            :::|||
 984 CTTCGATAACGACTGCAGTCTCT.........CCGAGTTGC......... 1015

 384 LysLeuHisHisCysProGlyAsnHisSerTrpAspSerThrIleSerGl 400
                 |||         |||||||||               ||
1016 ............TGT.........CACAGTTGG.........ACTCTGGG 1035

 400 ySerGlnArgAlaAlaPheCysAspHisLysThrThrProCysSerSerA 417
     ||||   |||                        :::|||   |||||||
1036 ATCTCCCAGACTCTTG...............AAGGCCCCGAAGAGCTCAG 1070

 417 laIleIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeuGlnPro 433
     ||   :::   ||||||            |||::::::      :::|||
1071 CCGAAGTT...CCTCTG............AATGCAAGCTCCCATCGTCCA 1105

 434 GlyIleAlaGln..................GlnTrpIleGlnSerLysAr 444
     ::::::|||:::                  ||||||   |||
1106 GCAGTGGCAAGAGGCTCTCGGGGGTGTCCTCAGTGGACTCAGCCTTTTCC 1155

 444 gGluAspIleValAsnGlnMetThrGluAlaCysLeuAsnGlnSerLeuA 461
         :::                  :::   |||
1156 TCCAGAG..................GATCGCTGT...............C 1172

 461 spAlaLeuLeuSerArgAspLeuIleMetLysGluAspTyrGluLeuVal 477
       :::|||||||||
1173 ACTGTCTTTTGAGC.................................... 1186

 478 SerThrLysProThrArgThrSerLysValArgGlnLeu..LeuAspThr 493
                          :::|||::::::|||     |||
1187 .....................GGGAAGCTTCAACAGGCGACCTGGGCCCC 1215

 494 ThrAspIleGlnGlyGluGluPheAlaLysValIleValGlnLysLeuLy 510
     ||||||||||||   ::::::::::::   :::|||::::::
1216 ACAGACATCCAGAAGAAGAAGCTAGTGGATGCCATCATATCAGGGGACAC 1265

 510 sAspAsnLysGlnMetGlyLeuGlnProTyrProGluIleLeuValValS 527
      ::::::   :::      |||||||||            ||||||::::
1266 CAGCAGGCTGATGAAGATCCTACAGCCCCAAGATGTGGACTTGGTTCTAG 1315

 527 erArgSerProSerLeu 532
     ::   |||   ||||||
1316 ACAGCAGTGCCAGCCTG 1332

seq_name: /cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-188-930-66

seq_documentation_block:
; Sequence 66, Application US/09188930A
```

```
; Patent No. 6150502
; GENERAL INFORMATION:
;  APPLICANT: Watson, James D.
;  APPLICANT: Strachan, Lorna
;  APPLICANT: Sleeman, Matthew
;  APPLICANT: Onrust, Rene
;  APPLICANT: Murison, James Greg
;  TITLE OF INVENTION: Compositions Isolated From Skin Cells
;  TITLE OF INVENTION: and Methods For Their Use
;  FILE REFERENCE: 11000.1011c1
;  CURRENT APPLICATION NUMBER: US/09/188,930A
;  CURRENT FILING DATE: 1998-11-09
;  NUMBER OF SEQ ID NOS: 348
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 66
;   LENGTH: 1888
;   TYPE: DNA
;   ORGANISM: mouse
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1690)...(1690)
;   NAME/KEY: unsure
;   LOCATION: (1755)...(1755)
;   NAME/KEY: unsure
;   LOCATION: (1864)...(1864)
US-09-188-930-66


alignment_scores:
           Quality:  548.00             Length:    541
             Ratio:  1.845                Gaps:     27
 Percent Similarity:  54.898   Percent Identity: 33.272

alignment_block:
 US-09-445-223-1 x US-09-188-930-66   ..

 Align seg 1/1  to: US-09-188-930-66  from: 1  to: 1888

      10 LeuProThrIleProTyrHisLysLeuAlaAspLeuArgTyrLeuSerAr 26
         |||   |||:::         :::::::|||     :::   ::::::
      43 CTGCGCACCTTCGACGCCGGCGAATTCGCAGGCTGGGAGAAGGTGGGCTC 92

      26 gGlyAlaSerGlyThrValSerSerAlaArgHisAlaAspTrpArgValG 43
          |||:::   |||   |||   :::::|||||||:::   |||::::::
      93 GGGCGGCTTCGGGCAGGTGTACAAGGTGCGCCATGTGCACTGGAAGACGT 142

      43 lnValAlaValLys.........HisLeuHisIleHisThrProLeuLeu 56
           :::|||:::|||            ||||||:::
     143 GGCTCGCGATCAAGTGCTCGCCCAGTCTGCACGTCGAC............ 180

      57 AspSerGluArgLysAspValLeuArgGluAlaGluIleLeuHisLysAl 73
         |||   ||||||   ::::::|||:::||||||:::   ::::::   ||
     181 GACAGGGAACGAATGGAGCTCCTGGAGGAAGCTAAGAAGATGGAGATGGC 230

      73 aArgPheSerTyrIlePheProIleLeuGlyIleCysAsnGluProGluP 90
         |:::|||   ||||||:::|||:::   |||||||||:::||||||
     231 CAAGTTCCGATACATTCTACCTGTGTACGGCATATGCCAGGAACCT.... 276

      90 heLeuGlyIleValThrGluTyrMetProAsnGlySerLeuAsnGluLeu 106
           :::|||:::|||   |||||||||   :::|||||||||::::::|||
     277 ..GTCGGCTTGGTCATGGAGTACATGGAGACAGGCTCCCTGGAGAAGCTG 324

     107 LeuHisArgLysThrGluTyrProAspValAlaTrpProLeuArgPheAr 123
         |||      :::               :::   |||   |||||||||||
     325 CTGGCCTCAGAG............CCATTGCCTTGGGACCTGCGCTTTCG 362

     123 gIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisAsnMetThrP 140
         ||||:::||||||   |||:::|||:::|||:::||||||   |||:::|
     363 CATCGTGCACGAGACAGCCGTGGGCATGAACTTCCTGCATTGCATGTCTC 412

     140 roProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeuAspAsn 156
         ||||||||||||||   |||||||||      |||||||||||||||
     413 CGCCACTGCTGCACCTAGACCTGAAGCCAGCGAACATCTTGCTGGATGCC 462

     157 GluPheHisVal...LysIleAlaAspPheGlyLeuSerLysTrpArgMe 172
         ::::::::::::   ::::::   ||||||||||||:::|||   :::
     463 CACTACCAAATGTCAAGATTTCTTGACTTTGGGCTGGCCAAGTGCAATGG 512

     172 tMetSerLeuSerGlnSerArgSerSerLysSerAlaProGluGly.... 187
          ||||||   |||::::::   |||               :::|||
     513 CATGTCCCACTCTCATGACCTCAGCATG...........GATGGCCTGT 550

     188 ..GlyThrIleIleTyrMetProProGluAsnTyrGluProGlyGlnLys 203
           |||||||||   |||:::||||||||||:::   :::      :::|||
     551 TTGGTACAATCGGCTACCTCCCTCCAGAGCGAATTCGT......GAGAAG 594

     204 SerArgAla...SerIleLysHisAspIleTyrSerTyrAlaValIleTh 219
         ||||||      :::   |||||||||:::||||||:::|||::::::
     595 AGCCGCTTGTTTGACACCAAACATGATGTATACAGCTTCGCCATTGTGAT 644

     219 rTrpGluValLeuSerArgLysGlnProPheGluAspValThrAsnProL 236
          |||   ||||||::::::::::::||||||   |||      |||   |
     645 CTGGGGTGTGCTTACACAGAATAATCCATTTGCAGATGAAAAGAACATCC 694

     236 euGlnIleMetTyrSerValSerGlnGlyHisArgProValIleAsnGlu 252
         ||:::||||||   :::|||   :::||||||||||||
     695 TACACATCATGATGAAAGTGGTAAAGGGCCACCGCCCA............ 732

     253 GluSerLeuProTyrAspIleProHisArg.........AlaArgMetIl 266
         |||      |||         |||:::            |||   :::||
     733 GAGCTGCCACCCATCTGCAGACCCCGGCCGCGTGCCTGTGCCAGCCTGAT 782

     266 eSerLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerP 283
         |:::|||::::::      |||      :::|||:::   ||||||
     783 AGGGCTCATGCAACGGTGCTGGCATGCAGACCCACAGGTGCGGCCC.... 828

     283 heLeuLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGlu 299
                                               ||||||:::|||
     829 ......................................ACCTTCCAAGAA 840

     300 IleThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSe 316
         ||||||
     841 ATTACC............................................ 846

     316 rValSerSerAlaIleHisLeuCysAspLysLysLysMetGluLeuSerL 333
             |||::::::      ||||||:::|||         |||::::::
     847 ....TCTGAAACAGAAGACCTTTGTGAGAAGCCTGATGAGGAGGTGAAA. 891

     333 euAsnIleProValAsnHisGlyProGlnGluGluSerCysGlySerSer 349
                    :::   |||   |||   |||:::|||      :::|||
     892 ........GACCTGGCTCATGAGCCAGGCGAGAAAAGCTCTCTAGAGTCC 933

     350 GlnLeuHisGluAsnSerGlySerProGluThrSerArgSerLeuProAl 366
         :::   :::   :::      |||   :::   :::|||:::
     934 AAGAGTGAGGCCAGGCCCGAGTCCTCACGCCTCAAGCGCGCCTCTGCTCC 983

     366 aProGlnAspAsnAsp.PheLeuSerArgLysAlaGlnAspCysTyrPhe 382
          |||   |||||||||    :::|||            :::|||
     984 CCCCTTCGATAACGACTGCAGTCTCT.........CCGAGTTGC...... 1018

     383 MetLysLeuHisHisCysProGlyAsnHisSerTrpAspSerThrIleSe 399
                        |||         |||||||||      |||:::::
    1019 ...............TGT.........CACAGTTGG......ACTCTGGG 1038

     399 rGlySerGlnArgAlaAlaPheCysAspHisLysThrThrProCysSerS 416
         ::::||||||               |||         :::|||   ||||
    1039 ATCTTCCCAA...............GACTCTTGAAAGGCCCCGAAGAGCT 1073

     416 erAlaIleIleAsnProLeuSerThrAlaGlyAsnSerGluArgLeuGln 432
         |||||   :::   ||||||            |||::::::      :::
    1074 CAGCCGAAGTT...CCTCTG............AATGCAAGCTCCCATCGT 1108

     433 ProGlyIleAlaGln..................GlnTrpIleGlnSerLy 443
         |||::::::|||:::                  ||||||   |||
    1109 CCAGCAGTGGCAAGAGGCTCTCGGGGGTGTCCTCAGTGGACTCAGCCTTT 1158
```

```
   443 sArgGluAspIleValAsnGlnMetThrGluAlaCysLeuAsnGlnSerL 460
              :::                  :::   |||
  1159 TCCTCCAGAG..................GATCGCTGT............. 1177

   460 euAspAlaLeuLeuSerArgAspLeuIleMetLysGluAspTyrGluLeu 476
            :::|||||||||
  1178 ..CACTGTCTTTTGAGC................................. 1192

   477 ValSerThrLysProThrArgThrSerLysValArgGlnLeu..LeuAsp 492
                               :::|||::::::|||     |||
  1193 ........................GGGAAGCTTCAACAGGCGACCTGGGC 1218

   493 ThrThrAspIleGlnGlyGluGluPheAlaLysValIleValGlnLysLe 509
          ||||||||||||   ::::::::::::   :::|||::::::
  1219 CCCACAGACATCCAGAAGAAGAAGCTAGTGGATGCCATCATATCAGGGGA 1268

   509 uLysAspAsnLysGlnMetGlyLeuGlnProTyrProGluIleLeuValV 526
           ::::::   :::      |||||||||            ||||||:
  1269 CACCAGCAGGCTGATGAAGATCCTACAGCCCCAAGATGTGGACTTGGTTC 1318

   526 alSerArgSerProSerLeu 532
       :::::   |||   ||||||
  1319 TAGACAGCAGTGCCAGCCTG 1338

seq_name: /cgn1_7/ptodata/1/ina/5A_COMB.seq:US-08-444-005-14

seq_documentation_block:
; Sequence 14, Application US/08444005
; Patent No. 5674734
;  GENERAL INFORMATION:
;    APPLICANT:  Leder, Philip
;    APPLICANT:  Seed, Brian
;    APPLICANT:  Stanger, Ben Z.
;    APPLICANT:  Lee, Tae-Ho
;    APPLICANT:  Kim, Emily
;    TITLE OF INVENTION:  CELL DEATH PROTEIN
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street, Suite 3100
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/444,005
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,164
;      REFERENCE/DOCKET NUMBER:  00383/026001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2268 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-444-005-14


alignment_scores:
            Quality:  397.00             Length:     609
                Ratio:  1.272              Gaps:      25
 Percent Similarity:  51.232   Percent Identity:  25.452

alignment_block:
 US-09-445-223-1 x US-08-444-005-14   ..

 Align seg 1/1  to: US-08-444-005-14  from: 1  to: 2268

    24 LeuSerArgGlyAlaSerGlyThrValSerSerAlaArgHisAlaAspTr 40
       |||:::   |||:::   |||   ||||||   :::   |||   :::
   118 CTAGACAGCGGAGGCTTCGGGAAGGTGTCCTTGTGTTACCACAGAAGCCA 167

    40 pArgValGlnValAlaValLysHisLeuHisIleHisThrProLeuLeuA 57
                 :::      |||         ::::::|||
   168 TGGATTTGTCATCCTGAAAAAA.........GTATACACAGGGCCCAACC 208

    57 spSerGluArgLysAspVal...LeuArgGluAlaGluIleLeuHisLys 72
         :::|||   ::::::|||   |||:::|||::::::::::::|||:::
   209 GCGCTGAGTACAATGAGGTTCTCTTGGAAGAGGGGAAGATGATGCACAGA 258

    73 AlaArgPheSerTyrIlePheProIleLeuGlyIleCysAsnGluProGl 89
          |||   |||   :::      :::|||||||||      |||
   259 CTGAGACACAGTCGAGTGGTGAAGCTACTGGGCATCATCATAGAAGAAGG 308

    89 uPheLeuGlyIleValThrGluTyrMetProAsnGlySerLeuAsnGluL 106
              ::::::|||   |||||||||   :::|||:::|||   ::::
   309 GAACTATTCGCTGGTGATGGAGTACATGGAGAAGGGCAACCTGATGCACG 358

   106 euLeuHisArgLysThrGluTyrProAspValAlaTrpProLeuArgPhe 122
       ::|||      ||||||:::      ||||||         |||:::
   359 TGCTA......AAGACCCAGATA...GATGTCCCACTTTCATTGAAAGGA 399

   123 ArgIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisAsnMetTh 139
       ||||||:::   |||         |||:::   |||||||||:::
   400 AGGATAATCGTGGAGGCCATAGAAGGCATGTGCTACTTACATGAC..... 444

   139 rProProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeuAspA 156
              ::::::|||   |||||||||   :::|||||||||:::|||:
   445 .AAAGGTGTGATACACAAGGACCTGAAGCCTGAGAATATCCTCGTTGATC 493

   156 snGluPheHisValLysIleAlaAspPheGlyLeuSerLysTrpArgMet 172
       :::::||||||:::||||||||||||:::|||:::::::::::::::
   494 GTGACTTTCACATTAAGATAGCCGATCTTGGTGTGGCTTCCTTTAAGACA 543

   173 MetSer...LeuSerGlnSerArgSerSerLysSerAlaPro........ 185
          |||   |||::::::::::::::::::|||:::
   544 TGGAGCAAACTGACTAAGGAGAAAGACAACAAGCAGAAAGAAGTGAGCAG 593

   186 ................GluGlyGlyThrIleIleTyrMetProProGlu. 196
                       :::|||||||||:::   ||||||  ||||||
   594 CACCACTAAGAAGAACAATGGTGGTACCCTTTACTACATGGCACCCGAAC 643

   197 ..............AsnTyrGluProGlyGlnLysSerArgAlaSerIle 208
                     |||   :::|||   :::||||||
   644 ACCTGAATGACATCAATGCAAAGCCCACGGAGAAGTCG............ 681

   209 LysHisAspIleTyrSerTyrAlaValIleThrTrpGluValLeuSerAr 225
             |||:::||||||::::::::::::   |||   :::::::::::
   682 ......GACGTGTACAGCTTTGGCATTGTCCTTTGGGCAATATTTGCAAA 725

   225 gLysGlnProPheGluAspValThrAsnProLeuGlnIleMetTyrSerV 242
       :|||:::|||:::|||:::|||            |||::::::      :
   726 AAAGGAGCCCTATGAGAATGTCATCTGTACTGAGCAGTTCGTGATCTGCA 775

   242 alSerGlnGlyHisArgProValIleAsnGlu......GluSerLeuPro 256
       ::::::::|||:::||||||   ::::::|||      |||      |||
   776 TAAAATCTGGGAACAGGCCAAATGTAGAGGAAATCCTTGAGTACTGTCCA 825

   257 TyrAspIleProHisArgAlaArgMetIleSerLeuIleGluSerGlyTr 273
          :::|||                  |||||||||:::|||      ||
   826 AGGGAGATC..................ATCAGCCTCATGGAGCGGTGCTG 857

   273 pAlaGlnAsnProAspGluArgProSerPheLeuLysCysLeuIleGluL 290
```

```
     |         |||::::::||||||:::||||||            |||:
 858 GCAGGCGATCCCAGAAGACAGGCCAACATTTCTTGGCATTGAAGAAGAAT 907

 290 euGluPro...ValLeuArgThrPheGluGluIleThrPheLeuGluAla 305
     :::::|||      |||      |||||||||   :::      |||
 908 TTAGGCCTTTTTACTTAAGTCATTTTGAAGAATATGTA...GAAGAGGAT 954

 306 ValIleGlnLeuLysLys..................ThrLysLeuGlnSe 316
     |||   :::|||||||||                        ||||||
 955 GTGGCAAGTTTAAAGAAAGAGTATCCAGATCAAAGCCCAGTGCTGCAGAG 1004

 316 rValSerSerAlaIleHisLeuCys......................... 324
      :::   |||      |||   |||
1005 AATGTTTTCACTGCAGCATGACTGTGTACCCTTACCTCCGAGCAGGTCAA 1054

 325 ..AspLysLysLysMetGluLeuSerLeuAsnIleProValAsnHisGly 340
       :::::::::      :::|||      :::      ::::::   |||
1055 ATTCAGAACAACCTGGATCGCTGCACAGTTCCCAGGGGCTCCAGATGGGT 1104

 341 ProGlnGluGluSerCysGlySerSerGln............LeuHisGl 353
     |||   |||||||||      ||||||:::                  ||
1105 CCTGTGGAGGAGTCCTGGTTTTCTTCCTCCCCAGAGTACCCACAGGACGA 1154

 353 uAsnSerGlySerProGluThr............................ 360
     ||||:::   |||   ::::::
1155 GAATGATCGCAGTGTGCAGGCTAAGCTGCAAGAGGAAGCCAGCTATCATG 1204

 361 ....................SerArgSerLeuProAlaProGlnAspAsn 370
                         :::::::::   |||   |||::::::|||
1205 CTTTTGGAATATTTGCAGAGAAACAGACAAAACCGCAGCCAAGGCAGAAT 1254

 371 AspPheLeuSerArgLysAlaGlnAspCysTyrPheMetLysLeuHisHi 387
     :::      :::|||:::   :::               ::::::   ||
1255 GAGGCTTACAACAGAGAGGAGGAA......AGGAAACGAAGGGTCTCTCA 1298

 387 sCysPro........................................... 389
     |   |||
1299 TGACCCCTTTGCACAGCAGAGAGCTCGTGAGAATATTAAGAGTGCAGGAG 1348

 390 ..GlyAsnHisSerTrpAspSerThrIleSerGlySerGlnArgAlaAla 405
          :::||||||      ||||||   |||         |||:::
1349 CAAGAGGTCATTCTGATCCCAGCACAACGAGT.........CGTGGAATT 1389

 406 PheCysAspHisLysThrThrProCysSerSerAlaIleIleAsn..... 420
           ::::::   :::   |||:::::::::::::::   |||
1390 GCAGTGCAACAGCTGTCATGGCCAGCCACCCAAACAGTTTGGAACAATGG 1439

 421 ................ProLeuSerThrAlaGlyAsnSerGluArgLeuG 432
                        ::::::|||:::|||::::::
1440 ATTGTATAATCAGCATGGATTTGGAACTACAGGTACAGGAGTTTGGTATC 1489

 432 lnProGlyIleAlaGlnGlnTrpIleGlnSerLys............... 443
       |||:::::::::|||::::::         |||
1490 CGCCAAATCTAAGCCAAATGTATAGTACTTATAAAACTCCAGTGCCAGAG 1539

 443 .................................................. 443

1540 ACCAACATACCGGGAAGCACACCCACCATGCCATACTTCTCTGGGCCAGT 1589

 444 .ArgGluAspIleVal.AsnGlnMet........................ 451
         :::|||:::::: |||   :::
1590 AGCAGATGACCTCATAAAATATACTATATTCAATAGTTCTGGTATTCAGA 1639

 452 ThrGluAlaCysLeuAsnGlnSerLeuAspAlaLeuLeuSerArgAspLe 468
        |||:::   :::         ||||||   :::   :::::::::
1640 TTGGAAACCACAATTATATGGATGTTGGACTGAATTCACAACCACCAAAC 1689

 468 uIleMetLysGluAspTyrGluLeuValSerThrLysProThrArgThrS 485
      |||:::   :::      :::|||   :::|||||||||:::      :
1690 AATACTTGCAAAGAAGAGTCGACTTCCAGACACCAAGCCATCTTTGATAA 1739

 485 erLysValArgGlnLeuLeuAspThrThrAspIleGlnGlyGlu...... 499
     ::   :::      |||::::::|||      ::::::|||:::
```

```
1740 CACCACTAGTCTGACTGATGAACACCTGAACCCTATCAGGGAAAACCTGG 1789

 500 ...........................GluPheAlaLysValIleValGl 507
                                ::::::|||::::::   :::::
1790 GAAGGCAGTGGAAAAACTGTGCCCGCAAGCTGGGCTTCACTGAGTCTCAG 1839

 507 nLysLeuLysAspAsnLysGlnMet 515
     :::::::|||::::::      |||
1840 ATCGATGAAATCGACCATGACTATG 1864
```

```
seq_name: /cgn1_7/ptodata/1/ina/6A_COMB.seq:US-09-329-418-2

seq_documentation_block:
; Sequence 2, Application US/09329418
; Patent No. 6096539
; GENERAL INFORMATION:
;  APPLICANT: ZENECA Limited
;  TITLE OF INVENTION: PROTEIN ACTIVATOR OF APOPTOSIS
;  FILE REFERENCE: PHM.70536
;  CURRENT APPLICATION NUMBER: US/09/329,418
;  CURRENT FILING DATE: 1999-06-11
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 1557
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-09-329-418-2


alignment_scores:
           Quality:  395.00            Length:    499
             Ratio:  1.452               Gaps:     21
 Percent Similarity:  54.509    Percent Identity: 28.056

alignment_block:
 US-09-445-223-1 x US-09-329-418-2   ..

 Align seg 1/1  to: US-09-329-418-2  from: 1  to: 1557

        11 ProThrIleProTyrHisLysLeuAlaAspLeuArgTyrLeuSerArgGl 27
           |||   :::      ::::::|||   :::   :::   :::::::::||
        40 CCCTTGGTGTCCATCGAGGAACTGGAGAACCAGGAGCTCGTCGGCAAAGG 89

        27 yAlaSerGlyThrValSerSerAlaArgHisAlaAspTrpArgValGlnV 44
           |:::   |||||||||      |||:::|||      |||      :::|
        90 CGGGTTCGGCACAGTGTTCCGGGCGCAACATAGGAAGTGGGGCTACGATG 139

        44 alAlaValLysHisLeuHisIleHisThrProLeuLeuAspSerGluArg 60
           ||||||||||||  ::::::
       140 TGGCGGTCAAGATCGTAAAC...........................TCG 162

        61 LysAspValLeuArgGluAlaGluIleLeuHisLysAlaArgPheSerTy 77
           |||   :::   ||||||::::::   :::   :::         :::::
       163 AAGGCGATATCCAGGGAGGTCAAGGCCATGGCAAGTCTGGATAACGAATT 212

        77 rIlePheProIleLeuGlyIleCysAsnGluProGluPhe.......... 90
           :::::::   :::   |||:::   ::::::   ::::::
       213 CGTGCTGCGCCTAGAAGGGGTTATCGAGAAGGTGAACTGGGACCAAGATC 262

        91 .....LeuGlyIleValThrGluTyrMetProAsnGlySerLeuAsnGlu 105
                   ::::::||||||::::::|||   ||||||||||||:::
       263 CCAAGCCGGCTCTGGTGACTAAATTCATGGAGAACGGCTCCTTGTCGGGG 312

       106 LeuLeuHisArgLysThrGluTyrProAspValAlaTrpProLeuArgPh 122
           ||||||:::   :::         |||         |||||||||
       313 CTGCTGCAGTCCCAGTGCCCTYGGCCC.........TGGCCGCTCCTTTG 353

       122 eArgIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisAsnMetT 139
            |||:::|||   |||::::::||||||:::   |||||||||::::::
       354 CCGCCTGCTGAAAGAAGTGGTGCTTGGGATGTTTTACCTGCACGACCAGA 403

       139 hrProProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeuAsp 155
```

```
     ::|||   |||||||||:::|||||||||   :::|||:::|||||||||
 404 ACCCGGTGCTCCTGCACCGGGACCTCAAGCCATCCAACGTCCTGCTGGAC 453

 156 AsnGluPheHisValLysIleAlaAspPheGlyLeuSerLysTrpArgMe 172
        |||:::|||||||||:::||||||||||||||||||   ::::::
 454 CCAGAGCTGCACGTCAAGCTGGCAGATTTTGGCCTGTCCACATTTCAG.. 501

 172 tMetSerLeuSerGlnSerArgSerSerLysSerAlaProGluGlyGlyT 189
         :::   |||||||||   ::::::::::::   |||   |||||||
 502 ....GGAGGCTCACAGTCAGGGACAGGGTCCGGGGAGCCA...GGGGGCA 544

 189 hrIleIleTyrMetProProGluAsnTyrGluProGlyGlnLysSerArg 205
     ||:::   |||:::   ||||||   :::      :::   :::   :::
 545 CCCTGGGCTACTTGGCCCCAGAACTGTTT...GTTAACGTAAACCGGAAG 591

 206 AlaSerIleLysHisAspIleTyrSerTyrAlaValIleThrTrpGluVa 222
     ||||||         |||:::||||||::::::::::::   |||   ||
 592 GCCTCCACAGCCAGTGACGTCTACAGCTTCGGGATCCTAATGTGGGCAGT 641

 222 lLeuSerArgLysGlnProPheGluAspValThrAsnProLeuGlnIleM 239
     ||||:::   ::::::      |||      |||:::|||      :::
 642 GCTTGCTGGAAGAGAA...GTTGAGTTGCCAACCGAACCATCACTCGTGT 688

 239 etTyrSerVal...SerGlnGlyHisArgProValIleAsnGluGluSer 254
          :::|||   ::::::   :::||||||   :::   |||
 689 ACGAAGCAGTGTGCAACAGGCAGAACCGGCCTTCATTGGCTGAG...... 732

 255 LeuProTyr.........AspIleProHisArgAlaArgMetIleSerLe 268
     ||||||            :::   |||            :::   :::||
 733 CTGCCCCAAGCCGGGCCTGAGACTCCCGGCTTAGAAGGACTGAAGGAGCT 782

 268 uIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerPheLeuL 285
     |::::::      |||:::::::::|||   :::||||||||||||   :
 783 AATGCAGCTCTGCTGGAGCAGTGAGCCCAAGGACAGACCCTCCTTCCAGG 832

 285 ysCysLeuIleGluLeuGluProValLeuArgThrPheGluGluIleThr 301
     ::||||||   :::   :::   |||::::::      |||      :::
 833 AATGCCTACCAAAAACTGATGAAGTCTTCCAGATGGTGGAG......AAC 876

 302 PheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSerValSe 318
        ::::::|||:::::::::      |||      |||:::::::::
 877 AATATGAATGCTGCTGTCTCCACGGTAAAGGATTTCCTGTCTCAGCTCAG 926

 318 rSer.AlaIle......HisLeuCysAspLysLysLysMetGlu...... 330
      ||| ||||||         |||   :::::::::      |||
 927 GAGCAGCAATAGGAGATTTTCTATCCCAGAGTCAGGCCAAGGAGGGACAG 976

 331 ......LeuSerLeuAsnIlePro.....................ValAs 338
           ::::::|||:::   |||                     :::
 977 AAATGGATGGCTTTAGGAGAACCATAGAAAACCAGCACTCTCGTAATGAT 1026

 338 n.HisGly...........ProGlnGluGluSerCysGlySerSerGln 350
       ||||||               |||   ||||||   |||:::||||||
1027 GTCATGGTTTCTGAGTGGCTAAACAAACTGAATCTAGAGGAGCCTCCCAG 1076

 351 LeuHisGluAsnSerGlySerProGluThrSerArgSerLeuProAlaPr 367
     |||                           |||   ::::::|||   ||
1077 CTC.......................TGTTCCTAAAAAATGCCCGAGCC 1102

 367 oGlnAspAsnAspPheLeuSerArgLysAla................... 377
     |   :::::::::      :::|||   |||
1103 TTACCAAGAGGAGCAGGGCACAAGAGGAGCAGGTTCCACAAGCCTGGACA 1152

 378 ............................................GlnAsp 379
                                                 :::|||
1153 GCAGGCACATCTTCAGATTCGATGGCCCAACCTCCCCAGACTCCAGAGAC 1202

 380 CysTyrPheMetLysLeuHisHisCysPro............GlyAsnHi 392
           |||:::|||            |||            ::::::
1203 CTCAACTTTCAGAAACCAGATGCCCAGCCCTACCTCAACTGGAACACCAA 1252

 392 sSerTrpAspSerThrIleSerGlySerGlnArgAlaAlaPheCysAspH 409
      ||||||            ||||||:::   :::
```

```
1253 GTCCTGGACCCCGAGGGAATCAGGGGGCTGAGAG................ 1286

 409 isLysThrThrProCysSerSerAlaIleIleAsnProLeuSerThrAla 425
             |||      :::   ::::::   :::|||   :::   :::
1287 ........ACAAGGCATGAACTGGTCCTGCAGGACCCCGGAGCCAAATCC 1328

 426 GlyAsnSerGluArgLeuGlnProGlyIleAlaGlnGlnTrpIleGlnSe 442
     :::|||            :::      |||         |||      ||
1329 AGTAACAGGGCGACCGCTCGTTAACATATACAACTGCTCTGGGGTGCAAG 1378

 442 rLysArgGluAspIleValAsnGlnMetThrGluAlaCysLeu 456
     |   |||:::::::::::::::      |||:::   ||||||
1379 TTGGAGACAACAACTACTTGACTATGCAACAGACAACTGCCTT 1421

seq_name: /cgn1_7/ptodata/1/ina/6A_COMB.seq:US-09-329-418-1

seq_documentation_block:
; Sequence 1, Application US/09329418
; Patent No. 6096539
; GENERAL INFORMATION:
;  APPLICANT: ZENECA Limited
;  TITLE OF INVENTION: PROTEIN ACTIVATOR OF APOPTOSIS
;  FILE REFERENCE: PHM.70536
;  CURRENT APPLICATION NUMBER: US/09/329,418
;  CURRENT FILING DATE: 1999-06-11
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 1873
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-09-329-418-1


alignment_scores:
           Quality:  392.50             Length:    501
             Ratio:  1.443                Gaps:     22
 Percent Similarity:  54.291   Percent Identity:  28.144

alignment_block:
 US-09-445-223-1 x US-09-329-418-1   ..

 Align seg 1/1  to: US-09-329-418-1  from: 1  to: 1873

      11 ProThrIleProTyrHisLysLeuAlaAspLeuArgTyrLeuSerArgGl 27
         |||   :::      ::::::|||   :::   :::   :::::::::||
     204 CCCTTGGTGTCCATCGAGGAACTGGAGAACCAGGAGCTCGTCGGCAAAGG 253

      27 yAlaSerGlyThrValSerSerAlaArgHisAlaAspTrpArgValGlnV 44
         |:::   |||||||||      |||:::|||      |||      :::|
     254 CGGGTTCGGCACAGTGTTCCGGGCGCAACATAGGAAGTGGGGCTACGATG 303

      44 alAlaValLysHisLeuHisIleHisThrProLeuLeuAspSerGluArg 60
         |||||||||||   ::::::
     304 TGGCGGTCAAGATCGTAAAC..........................TCG 326

      61 LysAspValLeuArgGluAlaGluIleLeuHisLysAlaArgPheSerTy 77
         |||   :::   ||||||::::::   :::   :::         :::::
     327 AAGGCGATATCCAGGGAGGTCAAGGCCATGGCAAGTCTGGATAACGAATT 376

      77 rIlePheProIleLeuGlyIle.....................CysAsnG 87
         :::::::   :::   |||:::                        ::::
     377 CGTGCTGCGCCTAGAAGGGGTTATCGAGAAGGTCGGCGGCTCGAGCCAAG 426

      87 luProGluPheLeuGlyIleValThrGluTyrMetProAsnGlySerLeu 103
         ::|||:::      ::::::||||||::::::|||   ||||||||||||
     427 ATCCCAAG...CCGGCTCTGGTGACTAAATTCATGGAGAACGGCTCCTTG 473

     104 AsnGluLeuLeuHisArgLysThrGluTyrProAspValAlaTrpProLe 120
         :::   ||||||:::   :::         |||         ||||||||
     474 TCGGGGCTGCTGCAGTCCCAGTGCCCTCGGCCC.........TGGCCGCT 514

     120 uArgPheArgIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisA 137
```

```
     |      |||:::|||   |||::::::||||||:::   |||||||||:
 515 CCTTTGCCGCCTGCTGAAAGAAGTGGTGCTTGGGATGTTTTACCTGCACG 564

 137 snMetThrProProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeu 153
     ::::::::|||   |||||||||:::|||||||||   :::|||:::|||
 565 ACCAGAACCCGGTGCTCCTGCACCGGGACCTCAAGCCATCCAACGTCCTG 614

 154 LeuAspAsnGluPheHisValLysIleAlaAspPheGlyLeuSerLysTr 170
     ||||||   |||:::||||||||||:::|||||||||||||||||   ::
 615 CTGGACCCAGAGCTGCACGTCAAGCTGGCAGATTTTGGCCTGTCCACATT 664

 170 pArgMetMetSerLeuSerGlnSerArgSerSerLysSerAlaProGluG 187
     ::::    . :::   |||||||||   ::::::::::::   |||   |
 665 TCAG......GGAGGCTCACAGTCAGGGACAGGGTCCGGGGAGCCA...G 705

 187 lyGlyThrIleIleTyrMetProProGluAsnTyrGluProGlyGlnLys 203
     ||||||||:::   |||:::   ||||||   :::      :::   :::
 706 GGGGCACCCTGGGCTACTTGGCCCCAGAACTGTTT...GTTAACGTAAAC 752

 204 SerArgAlaSerIleLysHisAspIleTyrSerTyrAlaValIleThrTr 220
        :::||||||         |||:::||||||::::::::::::   ||
 753 CGGAAGGCCTCCACAGCCAGTGACGTCTACAGCTTCGGGATCCTAATGTG 802

 220 pGluValLeuSerArgLysGlnProPheGluAspValThrAsnProLeuG 237
     |   ||||||:::   ::::::      |||      |||:::|||
 803 GGCAGTGCTTGCTGGAAGAGAA...GTTGAGTTGCCAACCGAACCATCAC 849

 237 lnIleMetTyrSerVal...SerGlnGlyHisArgProValIleAsnGlu 252
       :::      :::|||   ::::::   :::||||||   :::   |||
 850 TCGTGTACGAAGCAGTGTGCAACAGGCAGAACCGGCCTTCATTGGCTGAG 899

 253 GluSerLeuProTyr.........AspIleProHisArgAlaArgMetIl 266
           ||||||            :::   |||       .    :::
 900 ......CTGCCCCAAGCCGGGCCTGAGACTCCCGGCTTAGAAGGACTGAA 943

 266 eSerLeuIleGluSerGlyTrpAlaGlnAsnProAspGluArgProSerP 283
      :::|||::::::      |||:::::::::|||   :::||||||||||
 944 GGAGCTAATGCAGCTCTGCTGGAGCAGTGAGCCCAAGGACAGACCCTCCT 993

 283 heLeuLysCysLeuIleGluLeuGluProValLeuArgThrPheGluGlu 299
     ||   :::||||||   :::   :::   |||::::::      |||
 994 TCCAGGAATGCCTACCAAAAACTGATGAAGTCTTCCAGATGGTGGAG... 1040

 300 IleThrPheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSe 316
        :::   ::::::|||:::::::::      |||      |||:::::
1041 ...AACAATATGAATGCTGCTGTCTCCACGGTAAAGGATTTCCTGTCTCA 1087

 316 rValSerSer.AlaIle......HisLeuCysAspLysLysLysMetGlu 330
    .::::   ||| ||||||         |||   :::::::::      |||
1088 GCTCAGGAGCAGCAATAGGAGATTTTCTATCCCAGAGTCAGGCCAAGGAG 1137

 331 ............LeuSerLeuAsnIlePro.................... 336
                 ::::::|||:::   |||
1138 GGACAGAAATGGATGGCTTTAGGAGAACCATAGAAAACCAGCACTCTCGT 1187

 337 .ValAsn.HisGly............ProGlnGluGluSerCysGlySer 348
      :::    ||||||               |||   ||||||   |||:::
1188 AATGATGTCATGGTTTCTGAGTGGCTAAACAAACTGAATCTAGAGGAGCC 1237

 349 SerGlnLeuHisGluAsnSerGlySerProGluThrSerArgSerLeuPr 365
     |||||||||                           |||   ::::::||
1238 TCCCAGCTC...........................TGTTCCTAAAAAATGCC 1263

 365 oAlaProGlnAspAsnAspPheLeuSerArgLysAla............. 377
     |   |||   :::::::::      :::|||   |||
1264 CGAGCCTTACCAAGAGGAGCAGGGCACAAGAGGAGCAGGTTCCACAAGCC 1313

 377 .................................................. 377

1314 TGGACAGCAGGCACATCTTCAGATTCGATGGCCCAACCTCCCCAGACTCC 1363

 378 GlnAspCysTyrPheMetLysLeuHisHisCysPro............Gl 390
     :::|||      |||:::|||            |||            ::

1364 AGAGACCTCAACTTTCAGAAACCAGATGCCCAGCCCTACCTCAACTGGAA 1413

 390 yAsnHisSerTrpAspSerThrIleSerGlySerGlnArgAlaAlaPheC 407
     ::::   ||||||            ||||||:::   :::
1414 CACCAAGTCCTGGACCCCGAGGGAATCAGGGGGCTGAGAG.......... 1453

 407 ysAspHisLysThrThrProCysSerSerAlaIleIleAsnProLeuSer 423
                   |||      :::   ::::::   :::|||   :::
1454 ..............ACAAGGCATGAACTGGTCCTGCAGGACCCCGGAGCC 1489

 424 ThrAlaGlyAsnSerGluArgLeuGlnProGlyIleAlaGlnGlnTrpIl 440
        ::::::|||            :::      |||         |||
1490 AAATCCAGTAACAGGGCGACCGCTCGTTAACATATACAACTGCTCTGGGG 1539

 440 eGlnSerLysArgGluAspIleValAsnGlnMetThrGluAlaCysLeu 456
         |||   |||:::::::::::::::      |||:::   ||||||
1540 TGCAAGTTGGAGACAACAACTACTTGACTATGCAACAGACAACTGCCTT 1588

seq_name: /cgn1_7/ptodata/1/ina/5A_COMB.seq:US-07-928-464-1

seq_documentation_block:
; Sequence 1, Application US/07928464
; Patent No. 5367065
;  GENERAL INFORMATION:
;    APPLICANT:  Ecker, Joseph R.
;    APPLICANT:  Kieber, Joseph J.
;    TITLE OF INVENTION:  Constitutive Triple Response Gene and
;    TITLE OF INVENTION:  Mutations
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock, Washburn, Kurtz, Mackiewicz and
;      ADDRESSEE:  No. 5367065ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  U.S.A.
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/928,464
;      FILING DATE:  19920810
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Miller, Suzanne E.
;      REGISTRATION NUMBER:  32,279
;      REFERENCE/DOCKET NUMBER:  UPN-1086
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2890 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  118..2583
US-07-928-464-1


alignment_scores:
            Quality:  376.00           Length:    319
              Ratio:  1.861              Gaps:     12
 Percent Similarity:  63.323   Percent Identity:  30.721

alignment_block:
 US-09-445-223-1 x US-07-928-464-1   ..
```

```
Align seg 1/1  to: US-07-928-464-1  from: 1  to: 2890

   1 MetAsnGlyGluAlaIleCysSerAlaLeuProThr.............. 12
     ||||||:::      |||   :::   :::|||:::
1675 ATGAATGCCCCACCAATCAGTCAGCCAGTTCCAAACAGGGCAAATAGGGA 1724

  13 ............................IleProTyrHisLysLeuAlaA 20
                                 ||||||:::         :::|
1725 ACTTGGACTTGATGGTGATGATATGGACATCCCGTGG.........TGTG 1765

  20 spLeuArgTyrLeuSerArgGlyAlaSerGlyThrValSerSerAlaArg 36
     |||||:::      ::::::   ::::::|||:::   ::::::::::::
1766 ATCTTAATATAAAAGAAAAGATTGGAGCAGGTTCCTTTGGCACTGTCCAC 1815

  37 HisAlaAspTrpArg...ValGlnValAlaValLysHisLeu........ 49
     :::|||:::|||:::      :::||||||||||||   |||
1816 CGTGCTGAGTGGCATGGCTCGGATGTTGCTGTGAAAATTCTCATGGAGCA 1865

  50 .HisIleHisThrProLeuLeuAspSerGluArg...LysAspValLeuA 65
         :::|||               :::||||||   ::::::   ||||
1866 AGACTTCCAT...............GCTGAGCGTGTTAATGAGTTCTTAA 1900

  65 rgGluAlaGluIleLeuHisLysAlaArgPheSerTyrIlePheProIle 81
     |||||:::   |||:::   :::   |||         |||      :::
1901 GAGAGGTTGCGATAATGAAACGCCTTCGCCACCCTAACATTGTTCTCTTC 1950

  82 LeuGlyIleCysAsnGluProGluPheLeuGlyIleValThrGluTyrMe 98
     :::|||      ::::::|||      |||:::|||||||||||||||::
1951 ATGGGTGCGGTCACTCAACCTCCAAATTTGTCAATAGTGACAGAATATTT 2000

  98 tProAsnGlySerLeuAsnGluLeuLeuHisArgLysThrGluTyrProA 115
     :   :::|||||||||   :::|||||||||::::::            :
2001 GTCAAGAGGTAGTTTATACAGACTTTTGCATAAAAGTGGAGCAAGGGAGC 2050

 115 spValAlaTrpProLeuArgPheArgIleLeuHisGluIleAlaLeuGly 131
     :::::            |||:::   :::   :::::::::|||   |||
2051 AATTAGATGAGAGACGTCGCCTGAGTATGGCTTATGATGTGGCTAAGGGA 2100

 132 ValAsnTyrLeuHisAsnMetThrProProLeuLeuHisHisAspLeuLy 148
     :::|||||||||||||||   :::||||||::::::|||:::||||||||
2101 ATGAATTATCTTCACAATCGCAATCCTCCAATTGTGCATAGAGATCTAAA 2150

 148 sThrGlnAsnIleLeuLeuAspAsnGluPheHisValLysIleAlaAspP 165
     |:::   |||:::|||:::|||:::::::::   ||||||::::::||||
2151 ATCTCCAAACTTATTGGTTGACAAAAAATATACAGTCAAGGTTTGTGATT 2200

 165 heGlyLeuSerLysTrpArgMetMetSerLeuSerGlnSerArgSerSer 181
     |||||||||||:::   :::      ::::::   :::|||:::|||:::
2201 TTGGTCTCTCGCGATTGAAGGCCAGCACGTTTCTTTCCTCGAAGTCAGCA 2250

 182 LysSerAlaProGluGlyGlyThrIleIleTyrMetProProGlu..... 196
        ::::::||||||               :::|||   ||||||
2251 GCTGGAACCCCCGAG...............TGGATGGCACCAGAAGTCCT 2285

 197 .AsnTyrGluProGlyGlnLysSerArgAlaSerIleLysHisAspIleT 213
      :::   ||||||:::::::::               |||   |||:::|
2286 GCGAGATGAGCCGTCTAATGAA...............AAGTCAGATGTGT 2320

 213 yrSerTyrAlaValIleThrTrpGluValLeuSerArgLysGlnProPhe 229
     |||||::::::||||||   ||||||:::   :::   :::||||||:::
2321 ACAGCTTCGGGGTCATCTTGTGGGAGCTTGCTACATTGCAACAACCATGG 2370

 230 GluAspValThrAsnProLeuGlnIleMetTyrSerValSerGlnGlyHi 246
        ::::::   ||||||   |||::::::   :::|||:::
2371 GGTAACTTA...AATCCGGCTCAGGTTGTAGCTGCGGTTGGTTTCAAGTG 2417

 246 sArgProValIleAsnGluGluSerLeuProTyrAspIleProHis.... 261
      :::   :::                        :::||||||:::
2418 TAAACGGCTG........................GAGATCCCGCGTAATC 2443

 262 ..ArgAlaArgMetIleSerLeuIleGluSerGlyTrpAlaGlnAsnPro 277
       :::   ::::::   ::::::||||||:::   |||:::::::::|||
2444 TGAATCCTCAGGTTGCAGCCATAATCGAGGGTTGTTGGACCAATGAGCCA 2493
```

```
 278 AspGluArgProSerPheLeuLysCysLeuIleGluLeuGluProValLe 294
        :::|||||||||||| .       :::      |||:::|||:::::
2494 TGGAAGCGTCCATCATTTGCAACTATAATGGACTTGCTAAGACCATTGAT 2543

 294 uArgThr 296
     :::::::
2544 CAAATCA 2550

seq_name: /cgnl_7/ptodata/1/ina/PCTUS_COMB.seq:PCT-US93-07347-1

seq_documentation_block:
; Sequence 1, Application PC/TUS9307347
;  GENERAL INFORMATION:
;    APPLICANT:  Ecker, Joseph R.
;    APPLICANT:  Kieber, Joseph J.
;    TITLE OF INVENTION:  Constitutive Triple Response Gene and
;    TITLE OF INVENTION:  Mutations
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock, Washburn, Kurtz, Mackiewicz and
;      ADDRESSEE:  Norris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  U.S.A.
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/07347
;      FILING DATE:  19930805
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Miller, Suzanne E.
;      REGISTRATION NUMBER:  32,279
;      REFERENCE/DOCKET NUMBER:  UPN-1086
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2890 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  118..2583
PCT-US93-07347-1


alignment_scores:
            Quality:  376.00             Length:    319
              Ratio:  1.861                Gaps:     12
 Percent Similarity:  63.323   Percent Identity: 30.721

alignment_block:
 US-09-445-223-1 x PCT-US93-07347-1   ..

 Align seg 1/1  to: PCT-US93-07347-1  from: 1  to: 2890

       1 MetAsnGlyGluAlaIleCysSerAlaLeuProThr.............. 12
         ||||||:::      |||   :::   :::|||:::
    1675 ATGAATGCCCCACCAATCAGTCAGCCAGTTCCAAACAGGGCAAATAGGGA 1724

      13 ............................IleProTyrHisLysLeuAlaA 20
                                     ||||||:::         :::|
    1725 ACTTGGACTTGATGGTGATGATATGGACATCCCGTGG.........TGTG 1765
```

```
   20 spLeuArgTyrLeuSerArgGlyAlaSerGlyThrValSerSerAlaArg 36
      |||||:::      ::::::   ::::::|||:::   ::::::::::::
 1766 ATCTTAATATAAAAGAAAAGATTGGAGCAGGTTCCTTTGGCACTGTCCAC 1815

   37 HisAlaAspTrpArg...ValGlnValAlaValLysHisLeu........ 49
      :::|||:::|||:::      :::||||||||||||   |||
 1816 CGTGCTGAGTGGCATGGCTCGGATGTTGCTGTGAAAATTCTCATGGAGCA 1865

   50 .HisIleHisThrProLeuLeuAspSerGluArg...LysAspValLeuA 65
          :::|||               :::||||||   ::::::   ||||
 1866 AGACTTCCAT...............GCTGAGCGTGTTAATGAGTTCTTAA 1900

   65 rgGluAlaGluIleLeuHisLysAlaArgPheSerTyrIlePheProIle 81
      |||||:::   |||:::   :::   |||         |||      :::
 1901 GAGAGGTTGCGATAATGAAACGCCTTCGCCACCCTAACATTGTTCTCTTC 1950

   82 LeuGlyIleCysAsnGluProGluPheLeuGlyIleValThrGluTyrMe 98
      :::|||      ::::::|||      |||:::|||||||||||||||::
 1951 ATGGGTGCGGTCACTCAACCTCCAAATTTGTCAATAGTGACAGAATATTT 2000

   98 tProAsnGlySerLeuAsnGluLeuLeuHisArgLysThrGluTyrProA 115
      :   :::|||||||||   :::|||||||||::::::            :
 2001 GTCAAGAGGTAGTTTATACAGACTTTTGCATAAAAGTGGAGCAAGGGAGC 2050

  115 spValAlaTrpProLeuArgPheArgIleLeuHisGluIleAlaLeuGly 131
      :::::            |||:::   :::   :::::::::|||   |||
 2051 AATTAGATGAGAGACGTCGCCTGAGTATGGCTTATGATGTGGCTAAGGGA 2100

  132 ValAsnTyrLeuHisAsnMetThrProProLeuLeuHisHisAspLeuLy 148
      :::|||||||||||||||   :::||||||::::::|||:::||||||||
 2101 ATGAATTATCTTCACAATCGCAATCCTCCAATTGTGCATAGAGATCTAAA 2150

  148 sThrGlnAsnIleLeuLeuAspAsnGluPheHisValLysIleAlaAspP 165
      |:::   |||:::|||:::|||:::::::::   ||||||::::::||||
 2151 ATCTCCAAACTTATTGGTTGACAAAAAATATACAGTCAAGGTTTGTGATT 2200

  165 heGlyLeuSerLysTrpArgMetMetSerLeuSerGlnSerArgSerSer 181
      |||||||||||:::   :::      ::::::   :::|||:::|||:::
 2201 TTGGTCTCTCGCGATTGAAGGCCAGCACGTTTCTTTCCTCGAAGTCAGCA 2250

  182 LysSerAlaProGluGlyGlyThrIleIleTyrMetProProGlu..... 196
         ::::::||||||               :::|||   ||||||
 2251 GCTGGAACCCCCGAG...............TGGATGGCACCAGAAGTCCT 2285

  197 .AsnTyrGluProGlyGlnLysSerArgAlaSerIleLysHisAspIleT 213
       :::   ||||||:::::::::               |||   |||:::|
 2286 GCGAGATGAGCCGTCTAATGAA...............AAGTCAGATGTGT 2320

  213 yrSerTyrAlaValIleThrTrpGluValLeuSerArgLysGlnProPhe 229
      |||||::::::||||||   ||||||:::   :::   :::||||||:::
 2321 ACAGCTTCGGGGTCATCTTGTGGGAGCTTGCTACATTGCAACAACCATGG 2370

  230 GluAspValThrAsnProLeuGlnIleMetTyrSerValSerGlnGlyHi 246
         ::::::   ||||||   |||::::::   :::|||:::
 2371 GGTAACTTA...AATCCGGCTCAGGTTGTAGCTGCGGTTGGTTTCAAGTG 2417

  246 sArgProValIleAsnGluGluSerLeuProTyrAspIleProHis.... 261
       :::   :::                        :::||||||:::
 2418 TAAACGGCTG........................GAGATCCCGCGTAATC 2443

  262 ..ArgAlaArgMetIleSerLeuIleGluSerGlyTrpAlaGlnAsnPro 277
        :::   ::::::   ::::::||||||:::   |||:::::::::|||
 2444 TGAATCCTCAGGTTGCAGCCATAATCGAGGGTTGTTGGACCAATGAGCCA 2493

  278 AspGluArgProSerPheLeuLysCysLeuIleGluLeuGluProValLe 294
         :::||||||||||||         :::      |||:::|||:::::
 2494 TGGAAGCGTCCATCATTTGCAACTATAATGGACTTGCTAAGACCATTGAT 2543

  294 uArgThr 296
      :::::::
 2544 CAAATCA 2550

seq_name: /cgn1_7/ptodata/1/ina/5A_COMB.seq:US-08-003-311B-1
```

```
seq_documentation_block:
; Sequence 1, Application US/08003311B
; Patent No. 5444166
;  GENERAL INFORMATION:
;    APPLICANT:  Ecker, Joseph R.
;    APPLICANT:  Kieber, Joseph J.
;    TITLE OF INVENTION:  Constitutive Triple Response Gene
;    TITLE OF INVENTION:  and Mutations
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz and
;      ADDRESSEE:  No. 5444166ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  U.S.A.
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/003,311B
;      FILING DATE:  January 12, 1993
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/928,464
;      FILING DATE:  August 10, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lori Y. Beardell
;      REGISTRATION NUMBER:  34,293
;      REFERENCE/DOCKET NUMBER:  UPN-1108
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3033 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  118..2583
US-08-003-311B-1


alignment_scores:
            Quality:  376.00            Length:    319 .
              Ratio:  1.861               Gaps:     12
 Percent Similarity:  63.323   Percent Identity: 30.721

alignment_block:
 US-09-445-223-1 x US-08-003-311B-1   ..

 Align seg 1/1  to: US-08-003-311B-1  from: 1  to: 3033

        1 MetAsnGlyGluAlaIleCysSerAlaLeuProThr.............. 12
          ||||||:::      |||   :::   :::|||:::
     1675 ATGAATGCCCCACCAATCAGTCAGCCAGTTCCAAACAGGGCAAATAGGGA 1724

       13 ............................IleProTyrHisLysLeuAlaA 20
                                      ||||||:::         :::|
     1725 ACTTGGACTTGATGGTGATGATATGGACATCCCGTGG.........TGTG 1765

       20 spLeuArgTyrLeuSerArgGlyAlaSerGlyThrValSerSerAlaArg 36
          |||||:::      ::::::   ::::::|||:::   ::::::::::::
     1766 ATCTTAATATAAAAGAAAAGATTGGAGCAGGTTCCTTTGGCACTGTCCAC 1815

       37 HisAlaAspTrpArg...ValGlnValAlaValLysHisLeu........ 49
```

```
       :::|||:::|||:::      :::||||||||||||   |||
  1816 CGTGCTGAGTGGCATGGCTCGGATGTTGCTGTGAAAATTCTCATGGAGCA 1865

    50 .HisIleHisThrProLeuLeuAspSerGluArg...LysAspValLeuA 65
           :::|||               :::||||||   ::::::   ||||
  1866 AGACTTCCAT...............GCTGAGCGTGTTAATGAGTTCTTAA 1900

    65 rgGluAlaGluIleLeuHisLysAlaArgPheSerTyrIlePheProIle 81
       |||||:::   |||:::   :::   |||         |||      :::
  1901 GAGAGGTTGCGATAATGAAACGCCTTCGCCACCCTAACATTGTTCTCTTC 1950

    82 LeuGlyIleCysAsnGluProGluPheLeuGlyIleValThrGluTyrMe 98
       :::|||      ::::::|||      |||:::|||||||||||||||::
  1951 ATGGGTGCGGTCACTCAACCTCCAAATTTGTCAATAGTGACAGAATATTT 2000

    98 tProAsnGlySerLeuAsnGluLeuLeuHisArgLysThrGluTyrProA 115
       :   :::|||||||||   :::|||||||||::::::            :
  2001 GTCAAGAGGTAGTTTATACAGACTTTTGCATAAAAGTGGAGCAAGGGAGC 2050

   115 spValAlaTrpProLeuArgPheArgIleLeuHisGluIleAlaLeuGly 131
       :::::            |||:::   :::   :::::::::|||   |||
  2051 AATTAGATGAGAGACGTCGCCTGAGTATGGCTTATGATGTGGCTAAGGGA 2100

   132 ValAsnTyrLeuHisAsnMetThrProProLeuLeuHisHisAspLeuLy 148
       :::|||||||||||||||   :::||||||::::::|||:::||||||||
  2101 ATGAATTATCTTCACAATCGCAATCCTCCAATTGTGCATAGAGATCTAAA 2150

   148 sThrGlnAsnIleLeuLeuAspAsnGluPheHisValLysIleAlaAspP 165
       |:::   |||:::|||:::|||:::::::::   ||||||::::::||||
  2151 ATCTCCAAACTTATTGGTTGACAAAAAATATACAGTCAAGGTTTGTGATT 2200

   165 heGlyLeuSerLysTrpArgMetMetSerLeuSerGlnSerArgSerSer 181
       |||||||||||:::   :::      ::::::   :::|||:::|||:::
  2201 TTGGTCTCTCGCGATTGAAGGCCAGCACGTTTCTTTCCTCGAAGTCAGCA 2250

   182 LysSerAlaProGluGlyGlyThrIleIleTyrMetProProGlu..... 196
          ::::::||||||               :::|||   ||||||
  2251 GCTGGAACCCCCGAG...............TGGATGGCACCAGAAGTCCT 2285

   197 .AsnTyrGluProGlyGlnLysSerArgAlaSerIleLysHisAspIleT 213
        :::   ||||||:::::::::               |||   |||:::|
  2286 GCGAGATGAGCCGTCTAATGAA...............AAGTCAGATGTGT 2320

   213 yrSerTyrAlaValIleThrTrpGluValLeuSerArgLysGlnProPhe 229
       |||||::::::||||||   ||||||:::   :::   :::||||||:::
  2321 ACAGCTTCGGGGTCATCTTGTGGGAGCTTGCTACATTGCAACAACCATGG 2370

   230 GluAspValThrAsnProLeuGlnIleMetTyrSerValSerGlnGlyHi 246
          ::::::   ||||||   |||::::::   :::|||:::
  2371 GGTAACTTA...AATCCGGCTCAGGTTGTAGCTGCGGTTGGTTTCAAGTG 2417

   246 sArgProValIleAsnGluGluSerLeuProTyrAspIleProHis.... 261
        :::   :::                        :::||||||:::
  2418 TAAACGGCTG........................GAGATCCCGCGTAATC 2443

   262 ..ArgAlaArgMetIleSerLeuIleGluSerGlyTrpAlaGlnAsnPro 277
         :::   ::::::   ::::::||||||:::   |||:::::::::|||
  2444 TGAATCCTCAGGTTGCAGCCATAATCGAGGGTTGTTGGACCAATGAGCCA 2493

   278 AspGluArgProSerPheLeuLysCysLeuIleGluLeuGluProValLe 294
          :::||||||||||||         :::      |||:::|||:::::
  2494 TGGAAGCGTCCATCATTTGCAACTATAATGGACTTGCTAAGACCATTGAT 2543

   294 uArgThr 296
       :::::::
  2544 CAAATCA 2550

seq_name: /cgn1_7/ptodata/1/ina/5A_COMB.seq:US-08-261-432-1

seq_documentation_block:
; Sequence 1, Application US/08261432
; Patent No. 5602322
;  GENERAL INFORMATION:
;    APPLICANT:  Ecker, Joseph R.
;    APPLICANT:  Kieber, Joseph J.
;    TITLE OF INVENTION:  Constitutive Triple Response Gene
;    TITLE OF INVENTION:  and Mutations
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz and
;      ADDRESSEE:  No. 5602322ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  U.S.A.
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/261,432
;      FILING DATE:  June 17, 1994
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/003,311
;      FILING DATE:  January 12, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lori Y. Beardell
;      REGISTRATION NUMBER:  34,293
;      REFERENCE/DOCKET NUMBER:  UPN-1864
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3033 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  118..2583
US-08-261-432-1


alignment_scores:
           Quality:  376.00             Length:    319
             Ratio:  1.861                Gaps:     12
 Percent Similarity:  63.323   Percent Identity: 30.721

alignment_block:
 US-09-445-223-1 x US-08-261-432-1   ..

 Align seg 1/1  to: US-08-261-432-1  from: 1  to: 3033

       1 MetAsnGlyGluAlaIleCysSerAlaLeuProThr.............. 12
         ||||||:::      |||   :::   :::|||:::
    1675 ATGAATGCCCCACCAATCAGTCAGCCAGTTCCAAACAGGGCAAATAGGGA 1724

      13 ............................IleProTyrHisLysLeuAlaA 20
                                     ||||||:::         :::|
    1725 ACTTGGACTTGATGGTGATGATATGGACATCCCGTGG.........TGTG 1765

      20 spLeuArgTyrLeuSerArgGlyAlaSerGlyThrValSerSerAlaArg 36
         |||||:::      ::::::   ::::::|||:::   ::::::::::::
    1766 ATCTTAATATAAAAGAAAAGATTGGAGCAGGTTCCTTTGGCACTGTCCAC 1815

      37 HisAlaAspTrpArg...ValGlnValAlaValLysHisLeu........ 49
         :::|||:::|||:::      :::||||||||||||   |||
    1816 CGTGCTGAGTGGCATGGCTCGGATGTTGCTGTGAAAATTCTCATGGAGCA 1865

      50 .HisIleHisThrProLeuLeuAspSerGluArg...LysAspValLeuA 65
             :::|||               :::||||||   ::::::   ||||
```

```
1866 AGACTTCCAT...............GCTGAGCGTGTTAATGAGTTCTTAA 1900

  65 rgGluAlaGluIleLeuHisLysAlaArgPheSerTyrIlePheProIle 81
     |||||:::   |||:::   :::   |||         |||      :::
1901 GAGAGGTTGCGATAATGAAACGCCTTCGCCACCCTAACATTGTTCTCTTC 1950

  82 LeuGlyIleCysAsnGluProGluPheLeuGlyIleValThrGluTyrMe 98
     :::|||      ::::::|||      |||:::|||||||||||||||::
1951 ATGGGTGCGGTCACTCAACCTCCAAATTTGTCAATAGTGACAGAATATTT 2000

  98 tProAsnGlySerLeuAsnGluLeuLeuHisArgLysThrGluTyrProA 115
     :   :::|||||||||   :::|||||||||::::::            :
2001 GTCAAGAGGTAGTTTATACAGACTTTTGCATAAAAGTGGAGCAAGGGAGC 2050

 115 spValAlaTrpProLeuArgPheArgIleLeuHisGluIleAlaLeuGly 131
     :::::            |||:::   :::   :::::::::|||   |||
2051 AATTAGATGAGAGACGTCGCCTGAGTATGGCTTATGATGTGGCTAAGGGA 2100

 132 ValAsnTyrLeuHisAsnMetThrProProLeuLeuHisHisAspLeuLy 148
     :::|||||||||||||||   :::||||||::::::|||:::||||||||
2101 ATGAATTATCTTCACAATCGCAATCCTCCAATTGTGCATAGAGATCTAAA 2150

 148 sThrGlnAsnIleLeuLeuAspAsnGluPheHisValLysIleAlaAspP 165
    ·|:::   |||:::|||:::|||:::::::::   ||||||::::::||||
2151 ATCTCCAAACTTATTGGTTGACAAAAAATATACAGTCAAGGTTTGTGATT 2200

 165 heGlyLeuSerLysTrpArgMetMetSerLeuSerGlnSerArgSerSer 181
     |||||||||||:::   :::      ::::::   :::|||:::|||:::
2201 TTGGTCTCTCGCGATTGAAGGCCAGCACGTTTCTTTCCTCGAAGTCAGCA 2250

 182 LysSerAlaProGluGlyGlyThrIleIleTyrMetProProGlu..... 196
        ::::::||||||               :::|||   ||||||
2251 GCTGGAACCCCCGAG...............TGGATGGCACCAGAAGTCCT 2285

 197 .AsnTyrGluProGlyGlnLysSerArgAlaSerIleLysHisAspIleT 213
      :::   ||||||:::::::::               |||   |||:::|
2286 GCGAGATGAGCCGTCTAATGAA...............AAGTCAGATGTGT 2320

 213 yrSerTyrAlaValIleThrTrpGluValLeuSerArgLysGlnProPhe 229
     |||||::::::||||||   ||||||:::   :::   :::||||||:::
2321 ACAGCTTCGGGGTCATCTTGTGGGAGCTTGCTACATTGCAACAACCATGG 2370

 230 GluAspValThrAsnProLeuGlnIleMetTyrSerValSerGlnGlyHi 246
        ::::::   ||||||   |||::::::   :::|||:::
2371 GGTAACTTA...AATCCGGCTCAGGTTGTAGCTGCGGTTGGTTTCAAGTG 2417

 246 sArgProValIleAsnGluGluSerLeuProTyrAspIleProHis.... 261
      :::   :::                        :::||||||:::
2418 TAAACGGCTG.......................GAGATCCCGCGTAATC 2443

 262 ..ArgAlaArgMetIleSerLeuIleGluSerGlyTrpAlaGlnAsnPro 277
       :::   ::::::   ::::::||||||:::   |||:::::::::|||
2444 TGAATCCTCAGGTTGCAGCCATAATCGAGGGTTGTTGGACCAATGAGCCA 2493

 278 AspGluArgProSerPheLeuLysCysLeuIleGluLeuGluProValLe 294
        :::||||||||||||         :::     |||:::|||::::::
2494 TGGAAGCGTCCATCATTTGCAACTATAATGGACTTGCTAAGACCATTGAT 2543

 294 uArgThr 296
     :::::::
2544 CAAATCA 2550

seq_name: /cgn1_7/ptodata/1/ina/6B_COMB.seq:US-09-132-118-1,

seq_documentation_block:
; Sequence 1, Application US/09132118
; Patent No. 6211337
;  GENERAL INFORMATION:
;     APPLICANT:  BAICHWAL, VIJAY R
;     APPLICANT:  HUANG, JIANING
;     APPLICANT:  HSU, HAILING
;     APPLICANT:  GOEDDEL, DAVID V
;     TITLE OF INVENTION:  RIP: NOVEL HUMAN PROTEIN INVOLVED IN
```

```
;     TITLE OF INVENTION:  TUMOR NECROSIS FACTOR SIGNAL TRANSDUCTION, AND SCREENING
;     TITLE OF INVENTION:  ASSAYS
;     NUMBER OF SEQUENCES:  2
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;       STREET:  75 DENISE DRIVE
;       CITY:  HILLSBOROUGH
;       STATE:  CALIFORNIA
;       COUNTRY:  USA
;       ZIP:  94010
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/132,118
;       FILING DATE:
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  OSMAN, RICHARD A.
;       REGISTRATION NUMBER:  36,627
;       REFERENCE/DOCKET NUMBER:  T95-006-1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (650) 343-4341
;       TELEFAX:  (650) 343-4342
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2016 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  1..2013
US-09-132-118-1


alignment_scores:
            Quality:  369.00              Length:    389
              Ratio:  1.716                 Gaps:     12
 Percent Similarity:  55.270    Percent Identity: 28.792

alignment_block:
 US-09-445-223-1 x US-09-132-118-1   ..

 Align seg 1/1  to: US-09-132-118-1  from: 1  to: 2016

       24 LeuSerArgGlyAlaSerGlyThrValSerSerAlaArgHisAlaAspTr 40
          |||:::   |||:::   |||   ||||||   :::   |||
       67 CTGGACAGCGGAGGCTTTGGGAAGGTGTCTCTGTGTTTCCAC........ 108

       40 pArgValGln...ValAlaValLysHisLeuHisIleHisThrProLeuL 56
           |||:::|||   :::   :::                     |||
      109 .AGAACCCAGGGACTCATGATCATGAAAACAGTGTACAAGGGGCCCAACT 157

       56 euAspSerGluArgLysAspValLeuArgGluAlaGluIleLeuHisLys 72
               ::::::::::::   :::|||:::||||||:::::::::::::::
      158 GCATTGAGCACAACGAGGCCCTCTTGGAGGAGGCGAAGATGATGAACAGA 207

       73 AlaArgPheSerTyrIlePheProIleLeuGlyIleCysAsnGluProGl 89
             |||   |||   :::      :::||||||:::      |||
      208 CTGAGACACAGCCGGGTGGTGAAGCTCCTGGGCGTCATCATAGAGGAAGG 257

       89 uPheLeuGlyIleValThrGluTyrMetProAsnGlySerLeuAsnGluL 106
                 ::::::|||   |||||||||   :::|||:::|||   ::::
      258 GAAGTACTCCCTGGTGATGGAGTACATGGAGAAGGGCAACCTGATGCACG 307

      106 euLeuHisArgLysThrGluTyrProAspValAlaTrpProLeuArgPhe 122
          ::|||      :::   :::   |||               ::::::
      308 TGCTGAAAGCCGAGATGAGTACTCCG.........CTTTCTGTAAAAGGA 348
```

```
 123 ArgIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisAsnMetTh 139
     ||||||:::   ||||||      |||:::   |||||||||:::
 349 AGGATAATTTTGGAAATCATTGAAGGAATGTGCTACTTACATGGA..... 393

 139 rProProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeuAspA 156
            ::::::|||   |||||||||   :::|||||||||:::||||
 394 .AAAGGCGTGATACACAAGGACCTGAAGCCTGAAAATATCCTTGTTGATA 442

 156 snGluPheHisValLysIleAlaAspPheGlyLeuSerLysTrpArgMet 172
     ||:::||||||:::||||||||||||:::||||||::::::::::::|||
 443 ATGACTTCCACATTAAGATCGCAGACCTCGGCCTTGCCTCCTTTAAGATG 492

 173 MetSerLeuSerGlnSerArgSerSerLysSerAlaProGlu........ 186
        |||      ::::::::::::   ::::::      |||
 493 TGGAGCAAACTGAATAATGAAGAGCACAATGAGCTGAGGGAAGTGGACGG 542

 187 ................GlyGlyThrIleIleTyrMetProProGlu.... 196
                     |||||||||:::   ||||||   ||||||
 543 CACCGCTAAGAAGAATGGCGGCACCCTCTACTACATGGCGCCCGAGCACC 592

 197 ...........AsnTyrGluProGlyGlnLysSerArgAlaSerIleLys 209
                |||   :::|||   :::||||||
 593 TGAATGACGTCAACGCAAAGCCCACAGAGAAGTCG............... 627

 210 HisAspIleTyrSerTyrAlaValIleThrTrpGluValLeuSerArgLy 226
        |||:::||||||:::||||||:::   |||   ::::::::::::||
 628 ...GATGTGTACAGCTTTGCTGTAGTACTCTGGGCGATATTTGCAAATAA 674

 226 sGlnProPheGluAspValThrAsnProLeuGlnIleMetTyrSerValS 243
     |:::|||:::|||::::::            |||::::::      ::::
 675 GGAGCCATATGAAAATGCTATCTGTGAGCAGCAGTTGATAATGTGCATAA 724

 243 erGlnGlyHisArgProValIleAsnGlu......GluSerLeuProTyr 257
     :::::|||:::||||||   :::::::::      |||      |||
 725 AATCTGGGAACAGGCCAGATGTGGATGACATCACTGAGTACTGCCCAAGA 774

 258 AspIleProHisArgAlaArgMetIleSerLeuIleGluSerGlyTrpAl 274
     :::|||                  |||||||||::::::      |||
 775 GAAATT..................ATCAGTCTCATGAAGCTCTGCTGGGA 806

 274 aGlnAsnProAspGluArgProSerPheLeuLysCysLeuIleGluLeuG 291
         ||||||:::   ||||||:::|||               :::::::
 807 AGCGAATCCGGAAGCTCGGCCGACATTTCCTGGCATTGAAGAAAAATTTA 856

 291 luProValLeuArgThrPheGluGluIleThrPheLeuGluAlaValIle 307
     ::|||         :::      |||   :::      |||   |||
 857 GGCCTTTTTATTTAAGTCAATTAGAAGAAAGTGTAGAAGAGGACGTGAAG 906

 308 GlnLeuLysLys...........................ThrLysLeuGl 315
     :::|||||||||                              ::::::||
 907 AGTTTAAAGAAAGAGTATTCAAACGAAAATGCAGTTGTGAAGAGAATGCA 956

 315 nSerVal........................SerSerAlaIleHisLeuC 324
     ||||:::                        ||||||      :::   :
 957 GTCTCTTCAACTTGATTGTGTGGCAGTACCTTCAAGCCGGTCAAATTCAG 1006

 324 ysAspLysLysLysMetGluLeuSerLeuAsnIleProValAsnHisGly 340
     ::   ::::::      :::|||      :::      ::::::   |||
1007 CCACAGAACAGCCTGGTTCACTGCACAGTTCCCAGGGACTTGGGATGGGT 1056

 341 ProGlnGluGluSerCysGlySerSerGlnLeuHisGluAsnSerGlySe 357
     |||   |||||||||      :::   :::|||::::::
1057 CCTGTGGAGGAGTCCTGGTTTGCTCCTTCCCTGGAGCAC........... 1095

 357 rProGluThrSerArgSerLeuProAlaProGlnAspAsnAspPheLeuS 374
                           |||      :::::::::      |||:
1096 ......................CCACAAGAAGAGAATGAGCCCAGCCTGC 1123

 374 erArgLysAlaGlnAsp 379
     ::   |||   ||||||
1124 AGAGTAAACTCCAAGAC 1140

seq_name: /cgn1_7/ptodata/1/ina/5A_COMB.seq:US-08-444-005-16
```

```
seq_documentation_block:
; Sequence 16, Application US/08444005
; Patent No. 5674734
;  GENERAL INFORMATION:
;    APPLICANT:  Leder, Philip
;    APPLICANT:  Seed, Brian
;    APPLICANT:  Stanger, Ben Z.
;    APPLICANT:  Lee, Tae-Ho
;    APPLICANT:  Kim, Emily
;    TITLE OF INVENTION:  CELL DEATH PROTEIN
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street, Suite 3100
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/444,005
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,164
;      REFERENCE/DOCKET NUMBER:  00383/026001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2137 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-444-005-16


alignment_scores:
            Quality:  369.00             Length:    390
              Ratio:  1.732                Gaps:     13
 Percent Similarity:  54.615   Percent Identity: 29.231

alignment_block:
 US-09-445-223-1 x US-08-444-005-16   ..

 Align seg 1/1  to: US-08-444-005-16  from: 1  to: 2137

      24 LeuSerArgGlyAlaSerGlyThrValSerSerAlaArgHisAlaAspTr 40
         |||:::   |||:::   |||   ||||||   :::   |||
      67 CTGGACAGCGGAGGCTTCGGGAAGGTGTCTCTGTGTTTCCAC........ 108

      40 pArgValGln...ValAlaValLysHisLeuHisIleHisThrProLeuL 56
          |||:::|||   :::   :::                     |||
     109 .AGAACCCAGGGACTCATGATCATGAAAACAGTGTACAAGGGGCCCAACT 157

      56 euAspSerGluArgLysAspValLeuArgGluAlaGluIleLeuHisLys 72
              ::::::::::::   :::|||:::||||||:::::::::::::::
     158 GCATTGAGCACAACGAGGCCCTCTTGGAGGAGGCGAAGATGATGAACAGA 207

      73 AlaArgPheSerTyrIlePheProIleLeuGlyIleCysAsnGluProGl 89
            |||   |||   :::      :::||||||:::      |||
     208 CTGAGACACAGCCGGGTGGTGAAGCTCCTGGGCGTCATCATAGAGGAAGG 257

      89 uPheLeuGlyIleValThrGluTyrMetProAsnGlySerLeuAsnGluL 106
```

```
            :::::::|||  |||||||||   :::|||:::|||   ::::
 258 GAAGTACTCCCTGGTGATGGAGTACATGGAGAAGGGCAACCTGATGCACG 307

 106 euLeuHisArgLysThrGluTyrProAspValAlaTrpProLeuArgPhe 122
     ::|||      :::   :::   |||               ::::::
 308 TGCTGAAAGCCGAGATGAGTACTCCG.........CTTTCTGTAAAAGGA 348

 123 ArgIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisAsnMetTh 139
     ||||||:::   ||||||      |||:::   |||||||||:::
 349 AGGATAATTTTGGAAATCATTGAAGGAATGTGCTACTTACATGGA..... 393

 139 rProProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeuAspA 156
            :::::::|||  |||||||||   :::|||||||||:::||||
 394 .AAAGGCGTGATACACAAGGACCTGAAGCCTGAAAATATCCTTGTTGATA 442

 156 snGluPheHisValLysIleAlaAspPheGlyLeuSerLysTrpArgMet 172
     ||:::|||||||:::||||||||||||:::||||||::::::::::::|||
 443 ATGACTTCCACATTAAGATCGCAGACCTCGGCCTTGCCTCCTTTAAGATG 492

 173 MetSerLeuSerGlnSerArgSerSerLysSerAlaProGlu........ 186
        |||      ::::::::::::   ::::::      |||
 493 TGGAGCAAACTGAATAATGAAGAGCACAATGAGCTGAGGGAAGTGGACGG 542

 187 ................GlyGlyThrIleIleTyrMetProProGlu.... 196
                     |||||||||:::   ||||||   ||||||
 543 CACCGCTAAGAAGAATGGCGGCACCCTCTACTACATGGCGCCCGAGCACC 592

 197 ...........AsnTyrGluProGlyGlnLysSerArgAlaSerIleLys 209
                |||   :::|||   :::||||||
 593 TGAATGACGTCAACGCAAAGCCCACAGAGAAGTCG............... 627

 210 HisAspIleTyrSerTyrAlaValIleThrTrpGluValLeuSerArgLy 226
        |||:::||||||:::||||||:::   |||   :::::::::::::||
 628 ...GATGTGTACAGCTTTGCTGTAGTACTCTGGGCGATATTTGCAAATAA 674

 226 sGlnProPheGluAspValThrAsnProLeuGlnIleMetTyrSerValS 243
     |:::|||:::|||::::::            |||::::::      ::::
 675 GGAGCCATATGAAAATGCTATCTGTGAGCAGCAGTTGATAATGTGCATAA 724

 243 erGlnGlyHisArgProValIleAsnGlu......GluSerLeuProTyr 257
     :::::|||:::||||||   :::::::::      |||      |||
 725 AATCTGGGAACAGGCCAGATGTGGATGACATCACTGAGTACTGCCCAAGA 774

 258 AspIleProHisArgAlaArgMetIleSerLeuIleGluSerGlyTrpAl 274
     :::|||                  |||||||||::::::      |||
 775 GAAATT..................ATCAGTCTCATGAAGCTCTGCTGGGA 806

 274 aGlnAsnProAspGluArgProSerPheLeuLysCysLeuIleGluLeuG 291
         ||||||:::   ||||||:::|||               ::::::::
 807 AGCGAATCCGGAAGCTCGGCCGACATTTCCTGGCATTGAAGAAAAATTTA 856

 291 luProValLeuArgThrPheGluGluIleThrPheLeuGluAlaValIle 307
     ::|||         :::      |||   :::      |||   |||
 857 GGCCTTTTTATTTAAGTCAATTAGAAGAAAGTGTAGAAGAGGACGTGAAG 906

 308 GlnLeuLysLys...........................ThrLysLeuGl 315
     :::|||||||||                              :::::::||
 907 AGTTTAAAGAAAGAGTATTCAAACGAAAATGCAGTTGTGAAGAGAATGCA 956

 315 nSerVal........................SerSerAlaIleHisLeuC 324
     ||||:::                        ||||||      :::   :
 957 GTCTCTTCAACTTGATTGTGTGGCAGTACCTTCAAGCCGGTCAAATTCAG 1006

 324 ysAspLysLysLysMetGluLeuSerLeuAsnIleProValAsnHisGly 340
     ::   ::::::      :::|||      :::      ::::::   |||
1007 CCACAGAACAGGCTGGTTCACTGCACAGTTCCCAGGGACTTGGGATGGGT 1056

 341 ProGlnGluGluSer...CysGlySerSerGlnLeuHisGluAsnSerGl 356
     |||   |||||||||      |||   |||      |||
1057 CCTGTGGAGGAGTCCTGGTTTGGTCCTTCCCTGGAGCAC........... 1095

 356 ySerProGluThrSerArgSerLeuProAlaProGlnAspAsnAspPheL 373
                              |||      :::::::::      |

1096 .........................CCACAAGAAGAGAATGAGCCCAGCC 1120

 373 euSerArgLysAlaGlnAsp 379
     ||:::   |||   ||||||
1121 TGCAGAGTAAACTCCAAGAC 1140

seq_name: /cgn1_7/ptodata/1/ina/6A_COMB.seq:US-09-161-443-1

seq_documentation_block:
; Sequence 1, Application US/09161443A
; Patent No. 6020198
; GENERAL INFORMATION:
;  APPLICANT: C. Frank Bennett
;  APPLICANT: Lex M. Cowsert
;  TITLE OF INVENTION: ANTISENSE MODULATION OF RIP-1 EXPRESSION
;  FILE REFERENCE: RTS-0011
;  CURRENT APPLICATION NUMBER: US/09/161,443A
;  CURRENT FILING DATE: 1998-09-25
;  NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
;   LENGTH: 2617
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(2016)
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 2141
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 2311
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 2452
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 2496
;   OTHER INFORMATION: unknown
US-09-161-443-1


alignment_scores:
           Quality:  369.00              Length:    389
             Ratio:  1.716                 Gaps:     12
 Percent Similarity:  55.270   Percent Identity: 28.792

alignment_block:
 US-09-445-223-1 x US-09-161-443-1   ..

 Align seg 1/1  to: US-09-161-443-1  from: 1  to: 2617

  24 LeuSerArgGlyAlaSerGlyThrValSerSerAlaArgHisAlaAspTr 40
     |||:::   |||:::   |||   ||||||   :::   |||
  67 CTGGACAGCGGAGGCTTTGGGAAGGTGTCTCTGTGTTTCCAC........ 108

  40 pArgValGln...ValAlaValLysHisLeuHisIleHisThrProLeuL 56
      |||:::|||   :::   :::                     |||
 109 .AGAACCCAGGGACTCATGATCATGAAAACAGTGTACAAGGGGCCCAACT 157

  56 euAspSerGluArgLysAspValLeuArgGluAlaGluIleLeuHisLys 72
          ::::::::::::   :::|||:::||||||:::::::::::::::
 158 GCATTGAGCACAACGAGGCCCTCTTGGAGGAGGCGAAGATGATGAACAGA 207

  73 AlaArgPheSerTyrIlePheProIleLeuGlyIleCysAsnGluProGl 89
        |||   |||   :::      :::||||||:::      |||
 208 CTGAGACACAGCCGGGTGGTGAAGCTCCTGGGCGTCATCATAGAGGAAGG 257

  89 uPheLeuGlyIleValThrGluTyrMetProAsnGlySerLeuAsnGluL 106
            :::::::|||  |||||||||   :::|||:::|||   ::::
```

```
 258 GAAGTACTCCCTGGTGATGGAGTACATGGAGAAGGGCAACCTGATGCACG 307

 106 euLeuHisArgLysThrGluTyrProAspValAlaTrpProLeuArgPhe 122
     ::|||      :::   :::   |||               ::::::
 308 TGCTGAAAGCCGAGATGAGTACTCCG.........CTTTCTGTAAAAGGA 348

 123 ArgIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisAsnMetTh 139
     ||||||:::   ||||||      |||:::   |||||||||:::
 349 AGGATAATTTTGGAAATCATTGAAGGAATGTGCTACTTACATGGA..... 393

 139 rProProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeuAspA 156
            ::::::|||   |||||||||   :::|||||||||:::||||
 394 .AAAGGCGTGATACACAAGGACCTGAAGCCTGAAAATATCCTTGTTGATA 442

 156 snGluPheHisValLysIleAlaAspPheGlyLeuSerLysTrpArgMet 172
     ||:::||||||:::||||||||||||:::||||||::::::::::::|||
 443 ATGACTTCCACATTAAGATCGCAGACCTCGGCCTTGCCTCCTTTAAGATG 492

 173 MetSerLeuSerGlnSerArgSerSerLysSerAlaProGlu........ 186
        |||      ::::::::::::   ::::::      |||
 493 TGGAGCAAACTGAATAATGAAGAGCACAATGAGCTGAGGGAAGTGGACGG 542

 187 ................GlyGlyThrIleIleTyrMetProProGlu.... 196
                     |||||||||:::   ||||||   ||||||
 543 CACCGCTAAGAAGAATGGCGGCACCCTCTACTACATGGCGCCCGAGCACC 592

 197 ...........AsnTyrGluProGlyGlnLysSerArgAlaSerIleLys 209
                |||   :::|||   :::||||||
 593 TGAATGACGTCAACGCAAAGCCCACAGAGAAGTCG............... 627

 210 HisAspIleTyrSerTyrAlaValIleThrTrpGluValLeuSerArgLy 226
        |||:::||||||:::||||||:::   |||   ::::::::::::||
 628 ...GATGTGTACAGCTTTGCTGTAGTACTCTGGGCGATATTTGCAAATAA 674

 226 sGlnProPheGluAspValThrAsnProLeuGlnIleMetTyrSerValS 243
     |:::|||:::|||::::::            |||::::::      ::::
 675 GGAGCCATATGAAAATGCTATCTGTGAGCAGCAGTTGATAATGTGCATAA 724

 243 erGlnGlyHisArgProValIleAsnGlu......GluSerLeuProTyr 257
     :::::|||:::||||||   :::::::::      |||      |||
 725 AATCTGGGAACAGGCCAGATGTGGATGACATCACTGAGTACTGCCCAAGA 774

 258 AspIleProHisArgAlaArgMetIleSerLeuIleGluSerGlyTrpAl 274
     :::|||                  |||||||||::::::      |||
 775 GAAATT..................ATCAGTCTCATGAAGCTCTGCTGGGA 806

 274 aGlnAsnProAspGluArgProSerPheLeuLysCysLeuIleGluLeuG 291
         ||||||:::   ||||||:::|||               :::::::
 807 AGCGAATCCGGAAGCTCGGCCGACATTTCCTGGCATTGAAGAAAAATTTA 856

 291 luProValLeuArgThrPheGluGluIleThrPheLeuGluAlaValIle 307
     ::|||         :::      |||   :::      |||   |||
 857 GGCCTTTTTATTTAAGTCAATTAGAAGAAAGTGTAGAAGAGGACGTGAAG 906

 308 GlnLeuLysLys...........................ThrLysLeuGl 315
     :::|||||||||                              ::::::||
 907 AGTTTAAAGAAAGAGTATTCAAACGAAAATGCAGTTGTGAAGAGAATGCA 956

 315 nSerVal........................SerSerAlaIleHisLeuC 324
     ||||:::                        ||||||      :::   :
 957 GTCTCTTCAACTTGATTGTGTGGCAGTACCTTCAAGCCGGTCAAATTCAG 1006

 324 ysAspLysLysLysMetGluLeuSerLeuAsnIleProValAsnHisGly 340
     ::   ::::::      :::|||      :::      ::::::   |||
1007 CCACAGAACAGCCTGGTTCACTGCACAGTTCCCAGGGACTTGGGATGGGT 1056

 341 ProGlnGluGluSerCysGlySerSerGlnLeuHisGluAsnSerGlySe 357
     |||   |||||||||      :::   :::|||::::::
1057 CCTGTGGAGGAGTCCTGGTTTGCTCCTTCCCTGGAGCAC........... 1095

 357 rProGluThrSerArgSerLeuProAlaProGlnAspAsnAspPheLeuS 374
                           |||      :::::::::      |||:
1096 ......................CCACAAGAAGAGAATGAGCCCAGCCTGC 1123
```

```
 374 erArgLysAlaGlnAsp 379
     ::   |||   ||||||
1124 AGAGTAAACTCCAAGAC 1140
```

seq_name: /cgn1_7/ptodata/1/ina/5A_COMB.seq:US-08-700-575-31

```
seq_documentation_block:
; Sequence 31, Application US/08700575
; Patent No. 5817479
;  GENERAL INFORMATION:
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Hawkins, Phillip R.
;    APPLICANT:  Wilde, Craig G.
;    TITLE OF INVENTION:  NOVEL HUMAN KINASE HOMOLOGS
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/700,575
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BILLINGS, LUCY J
;      REGISTRATION NUMBER:  36749
;      REFERENCE/DOCKET NUMBER:  SP-100 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  210 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      LIBRARY:  THP-1 Phorbol LPS
;      CLONE:  156108
US-08-700-575-31


alignment_scores:
           Quality:  367.00              Length:     69
             Ratio:  5.319                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-445-223-1 x US-08-700-575-31   ..

 Align seg 1/1  to: US-08-700-575-31  from: 1  to: 210

 196 GluAsnTyrGluProGlyGlnLysSerArgAlaSerIleLysHisAspIl 212
     ||||||||||||||||||||||||||||||||||||||||||||||||||
   2 GAAACTATGAACCTGGACAAAAATCAAGGGCCAGTATCAAGCACGATAT 51

 212 eTyrSerTyrAlaValIleThrTrpGluValLeuSerArgLysGlnProP 229
     ||||||||||||||||||||||||||||||||||||||||||||||||||
  52 ATATAGCTATGCAGTTATCACATGGAAGTGTTATCCAGAAAACAGCCTT 101

 229 heGluAspValThrAsnProLeuGlnIleMetTyrSerValSerGlnGly 245
     ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
     102 TTGAAGATGTCACCAATCCTTTGCAGATAATGTATAGTGTGTCACAAGGA 151

     246 HisArgProValIleAsnGluGluSerLeuProTyrAspIleProHisAr 262
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     152 CATCGACCTGTTATTAATGAAGAAAGTTTGCCATATGATATACCTCACCG 201

     262 gAlaArg 264
         |||||||
     202 AGCACGT 208

seq_name: /cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-685-625A-1

seq_documentation_block:
; Sequence 1, Application US/08685625A
; Patent No. 5945301
;  GENERAL INFORMATION:
;    APPLICANT:  UENO, Naoto
;    APPLICANT:  MATSUMOTO, Kunihiro
;    APPLICANT:  IRIE, Kenji
;    TITLE OF INVENTION:  NOVEL KINASE IN TGF-BETA FAMILY SIGNAL
;    TITLE OF INVENTION:  TRANSDUCTION SYSTEM
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS, L.L.P.
;      STREET:  P.O. Box 1404
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  United States
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/685,625A
;      FILING DATE:  24-JUL-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-253549
;      FILING DATE:  29-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meuth, Donna M.
;      REGISTRATION NUMBER:  36,607
;      REFERENCE/DOCKET NUMBER:  001560-267
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-6620
;      TELEFAX:  (703) 836-2021
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2443 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  157..1893
US-08-685-625A-1


alignment_scores:
            Quality:  315.50             Length:    589
              Ratio:  1.119                Gaps:     29
 Percent Similarity:  47.878   Percent Identity: 25.127

alignment_block:
 US-09-445-223-1 x US-08-685-625A-1   ..

 Align seg 1/1  to: US-08-685-625A-1  from: 1  to: 2443

      13 IleProTyrHisLysLeuAlaAspLeuArgTyrLeuSerArgGlyAlaSe 29
         |||   |||   ::::::         :::   ::::::|||||||||

 247 ATCGACTACAAGGAGATCGAGGTGGAAGAGGTTGTCGGAAGAGGAGCTTT 296

  29 rGlyThrValSerSerAlaArgHisAlaAspTrpArgValGln...ValA 45
      |||:::|||   :::|||:::         ||||||::::::   ||||
 297 TGGAGTAGTTTGCAAAGCTAAG.........TGGAGAGCAAAAGATGTCG 337

  45 laValLysHisLeuHisIleHisThrProLeuLeuAspSerGluArgLys 61
     ||:::|||:::::::::                  :::||||||||||||
 338 CTATTAAACAGATAGAA...............AGTGAGTCTGAGAGGAAG 372

  62 AspValLeuArgGluAlaGluIleLeuHisLysAlaArgPheSerTyrIl 78
           :::   |||   :::   |||   :::::::::         ||
 373 GCTTTCATTGTGGAGCTCCGGCAGTTGTCGCGTGTGAACCATCCTAACAT 422

  78 ePheProIleLeuGlyIleCysAsnGluProGluPheLeuGlyIleValT 95
     |      :::   |||   |||   :::|||      :::   :::|||
 423 TGTCAAGTTGTACGGAGCCTGCCTGAATCCA......GTATGTCTTGTGA 466

  95 hrGluTyrMetProAsnGlySerLeuAsnGluLeuLeuHisArgLysThr 111
       ||||||      :::|||||||||   ::::::||||||
 467 TGGAATATGCAGAGGGGGGCTCATTGTATAATGTGCTGCATGGTGCTGAA 516

 112 GluTyrPro..................AspValAlaTrpProLeuArgPh 122
           |||                     ::::::|||   |||
 517 CCATTGCCTTACTACACTGCTGCTCATGCCATGAGCTGGTGTTTA..... 561

 122 eArgIleLeuHisGluIleAlaLeuGlyValAsnTyrLeuHisAsnMetT 139
                  :::   :::   ||||||   |||||||||:::|||
 562 .............CAGTGTTCCCAAGGAGTGGCTTACCTGCACAGCATGC 598

 139 hrPro...ProLeuLeuHisHisAspLeuLysThrGlnAsnIleLeuLeu 154
       |||      |||:::|||:::|||||||||      |||:::||||||
 599 AGCCCAAAGCGCTGATTCACAGGGACCTCAAGCCTCCAAACTTGCTGCTG 648

 155 ...AspAsnGluPheHisValLysIleAlaAspPheGlyLeuSerLysTr 170
           :::         :::||||||:::|||||||||   :::
 649 GTTGCAGGAGGGACAGTTCTAAAAATCTGCGATTTTGGTACAGCTTGTGA 698

 170 pArgMetMetSerLeuSerGlnSerArgSerSerLysSerAlaProGluG 187
         :::      :::::::::::::::
 699 CATCCAAACACACATGACCAATAATAAA...................... 726

 187 lyGlyThrIleIleTyrMetProProGluAsnTyrGluProGlyGlnLys 203
       |||:::      :::|||   ||||||   :::|||
 727 ..GGGAGTGCTGCTTGGATGGCGCCTGAAGTGTTTGAA.........GGT 765

 204 SerArgAlaSerIleLysHisAspIleTyrSerTyrAlaValIleThrTr 220
     |||:::   |||   |||   |||::::::|||:::::::::|||   ||
 766 AGCAATTACAGTGAAAAGTGTGATGTCTTCAGCTGGGGTATTATCCTCTG 815

 220 pGluValLeuSerArgLysGlnProPheGluAspValThrAsnPro...L 236
     |||||||::::::|||::::::||||||:::::::::   :::|||   :
 816 GGAAGTGATAACACGCCGGAAACCCTTCGATGAGATCGGTGGCCCAGCTT 865

 236 euGlnIleMetTyrSerValSerGlnGlyHisArgProValIleAsnGlu 252
     :::::||||||::::::|||   :::|||   ||||||   :::
 866 TCAGAATCATGTGGGCTGTTCATAATGGCACTCGACCACCACTG...ATC 912

 253 GluSerLeuProTyrAspIleProHisArgAlaArgMetIleSerLeuIl 269
     ::::::||||||      |||                     ||||||::
 913 AAAAATTTACCTAAGCCCATTGAG..................AGCTTGAT 944

 269 eGluSerGlyTrpAlaGlnAsnProAspGluArgProSer.........P 283
     :         |||:::::::::|||::::::|||||||||         :
 945 GACACGCTGTTGGTCTAAGGACCCATCTCAGCGCCCTTCAATGGAGGAAA 994

 283 heLeuLysCysLeuIleGluLeu........................... 290
     :::::|||   :::   :::|||
 995 TTGTGAAAATAATGACTCACTTGATGCGGTACTTCCCAGGAGCGGATGAG 1044

 290 .................................................. 290

1045 CCATTACAGTATCCTTGTCACTACTCTGATGAAGGGCAGAGCAACTCAGC 1094
```

```
 290 .................................................. 290

1095 CACCAGCACAGGCTCGTTCATGGACATTGCTTCTACAAATACCAGTAATA 1144

 291 .................GluProValLeuArgThrPheGluGluIleThr 301
                      |||   |||      |||   :::   |||
1145 AAAGTGACACAAATATGGAACAGGTTCCTGCCACAAACGACACTATTAAA 1194

 302 PheLeuGluAlaValIleGlnLeuLysLysThrLysLeuGlnSerValSe 318
        ||||||:::   :::      :::::::::|||   ||||||   ||
1195 CGCTTGGAGTCAAAACTGTTGAAAAACCAGGCAAAGCAACAGAGTGAATC 1244

 318 rSerAlaIleHisLeuCysAspLysLysLysMetGluLeuSerLeuAsnI 335
     |:::   :::   |||      ::::::      :::::::::
1245 TGGACGCCTGAGCTTGGGAGCCTCTCGTGGGAGCAGTGTGGAGAGCTTGC 1294

 335 leProValAsnHisGlyProGlnGluGluSerCysGlySerSerGlnLeu 351
       |||:::::::::|||   ::::::   |||:::      |||::::::
1295 CCCCCACTTCCGAGGGC...AAGAGGATGAGTGCTGACATGTCTGAAATA 1341

 352 His..................GluAsnSerGlySerProGluThrSerAr 362
     :::                     ::::::|||:::|||:::      ||
1342 GAAGCCAGGATCGTGGCGACTGCAGGTAACGGGCAACCAAGG...CGTAG 1388

 362 gSerLeuProAlaProGlnAspAsnAspPheLeuSerArgLysAlaGlnA 379
     ||||:::         ||||||                           |
1389 ATCCATC.........CAAGACTT.........................G 1404

 379 spCysTyrPheMetLysLeuHisHisCysProGlyAsnHisSerTrpAsp 395
     ||||||||                                    ||||||
1405 ACTGTTAC....................................TGGGAC 1418

 396 SerThrIleSerGlySerGln............................. 402
        |||   ||||||:::|||
1419 AGAACCTGGTCAGGTGAGCAGCCGGTCATCCAGCCCTAGTGTCAGAATGA 1468

 403 ......................ArgAlaAlaPheCysAsp.HisLysThr 411
                           |||   |||   |||::: |||
1469 TCACTACCTCAGGACCAACCTCAGAGAAGCCAGCTGCGAGTCACCCATGG 1518

 412 ThrProCysSerSerAlaIleIleAsn..................ProLe 422
     ||||||   :::|||:::      |||                  |||::
1519 ACCCCTGATGATTCCACAGACACCAATGGCTCAGATAACTCCATCCCAAT 1568

 422 uSerThrAlaGlyAsnSerGluArgLeuGlnPro................ 433
     ::::            :::::::::|||||||||
1569 GGCGTATCTTACACTGGATCACCAGCTACAGCCTCTAGCGCCGTGCCCAA 1618

 434 .................GlyIleAlaGlnGlnTrpIleGlnSerLysArg 444
                      ::::::   :::|||      :::      :::
1619 ACTCCAAAGAATCCATGGCAGTGTTCGAACAGCACTGTAAAATGGCACAG 1668

 445 GluAspIleValAsnGlnMetThrGluAlaCysLeu.........AsnGl 458
     |||   :::      |||         |||   |||         :::||
1669 GAGTATATGAAAGTTCAAACCGAAATCGCATTGTTACTACAGAGAAAGCA 1718

 458 nSerLeuAspAlaLeuLeuSerArgAspLeuIleMetLysGluAspTyrG 475
     |:::|||   |||   |||::::::|||         ::::::|||   :
1719 AGAACTAGTTGCAGAATTGGACCAGGAT.........GAAAAGGACCAGC 1759

 475 luLeuValSerThrLysProThrArgThrSerLysVal............ 487
     ::                     :::||||||::::::
1760 AA.....................AATACATCTCGTCTGGTACAGGAACAT 1788

 488 ArgGlnLeuLeuAsp 492
     ::::::|||||||||
1789 AAAAAGCTTTTAGAT 1803
```

Access DB# ___________

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: ______________________ Examiner # : __________ Date: __________

Art Unit: __________ Phone Number 30__________ Serial Number: ____________________

Mail Box and Bldg/Room Location: ____________ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

****

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: ____

Inventors (please provide full names): ____

____

Earliest Priority Filing Date: ____________________

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

****

STAFF USE ONLY	**Type of Search**	**Vendors and cost where applicable**
Searcher: Toby Port	NA Sequence (#) 2	STN
Searcher Phone #: 308-3534	AA Sequence (#) 1	Dialog
Searcher Location:	Structure (#)	Questel/Orbit
Date Searcher Picked Up: 6/13	Bibliographic	Dr.Link
Date Completed: 6/14	Litigation	Lexis/Nexis
Searcher Prep & Review Time: 10	Fulltext	Sequence Systems g
Clerical Prep Time:	Patent Family	WWW/Internet
Online Time: 10	Other	Other (specify)

PTO-1590 (1-2000)